

研究速報

マルチメディア情報の補足による初心者向け料理レシピの作成へ向けて

志士地由香^{†*} 井手 一郎^{†a)} (正員: シニア会員)中村 裕一^{††} (正員) 出口 大輔[†] (正員)高橋 友和^{†††} (正員) 村瀬 洋[†] (正員: フェロー)

Towards the Creation of a Cooking Recipe for Inexperienced Users by Supplementing Multimedia Information

Yuka SHIDOCHI^{†*}, Nonmember,Ichiro IDE^{†a)}, Senior Member, Yuichi NAKAMURA^{††},Daisuke DEGUCHI^{††}, Tomokazu TAKAHASHI^{†††}, Members, and Hiroshi MURASE[†], Fellow[†]名古屋大学大学院情報科学研究科, 名古屋市

Graduate School of Information Science, Nagoya University, Nagoya-shi, 464-8601 Japan

^{††}京都大学学術情報メディアセンター, 京都市

Academic Center for Computing and Media Sciences, Kyoto University, Kyoto-shi, 606-8501 Japan

^{†††}岐阜聖徳学園大学経済情報学部, 岐阜市

Faculty of Economics and Information, Gifu Shotoku Gakuen University, Gifu-shi, 500-8288 Japan

* 現在, トヨタ自動車株式会社

a) E-mail: ide@is.nagoya-u.ac.jp

あらまし 既存のテキスト料理レシピから初心者に分かりにくい表現を検出した後, 大量の料理レシピから抽出した情報に基づいてテキストを, 調理動作の種類に応じて画像・映像を補足することで, 分かりやすいマルチメディア料理レシピを作る方法を検討した.

キーワード 料理レシピ, 書換え, マルチメディア化
1. ま え が き

近年, 情報通信技術による家事の支援が現実的になりつつある. 我々はその中でも, 台所における調理活動の支援に注目している. 調理活動は豊富な経験と知識を要するため, 初心者には必ずしも容易ではない.

このような問題意識のもとに, 浜田らは料理番組の副読本中の料理レシピについて調理手順構造を解析 [1] し, 調理手順ごとに映像と対応付け [2], 必要な映像を見ながら調理できる「Cooking Navi」インタフェース [3] を開発した. 更に近年, 商用のゲーム機向けにも調理支援ソフトウェアが発売された^(注1). しかし, これらは特定の料理レシピについてマルチメディア情報を選択的に閲覧できるようにすることで, 分かりやすく説明しているだけである.

一方, ウェブ上にテキストと画像による説明を中心とした料理レシピが増加しているが, その多くは専門家が作成したのではないため, 必ずしも初心者向けに分かりやすく記述されているとは限らない.

表 1 調理動作に応じた画像・映像の役割

Table 1 Roles of image and video contents according to cooking operations.

調理動作	混合	加熱	切削	装飾	浸漬	冷却	分離	その他	合計
画像	6	17	13	3	0	2	2	27	70
連続画像	0	0	30	0	0	1	0	5	36
映像	18	12	5	0	0	4	0	8	47
合計	24	29	48	3	0	7	2	40	153

我々はこれらの問題を踏まえ, 既存のテキスト料理レシピから初心者に分かりにくい表現を検出 [4] し, テキスト・画像・映像のマルチメディア情報により補足 [5] する研究に取り組んでいる [6]. 本論文ではこの研究について, これまで得られた検討結果を報告する.

利用者の要求に応じて料理レシピを書き換える研究として, 宮脇らは高次脳機能障害者を対象としてテキストを書き換える方法を提案した [7]. しかし, この方法は長い記述を短く分割するとともに省略語句を補足するものであり, 外部情報を用いることは想定していない. 一方, 外部情報を用いる方法として, 我々 [8] や一連の“Computer Cooking Contest”^(注2)参加者 [9], [10] は, 料理レシピ中の素材及び関連する調理手順を書き換える方法を提案したが, 調理者の経験や知識を考慮した方法ではない上に, マルチメディア情報による補足は想定していない.

2. 初心者向け料理レシピ固有の表現方法の分析

初心者向けの料理レシピを作成する際に考慮すべき点を明らかにするために, 専門家が作成した料理レシピを掲載しているウェブサイト^(注3)から取得したテキストと画像による説明からなる, 24 対の料理レシピを分析した. 各対は, 同一料理に関する子供向けと一般向けの料理レシピからなる. ここでは, 子供向け料理レシピを初心者向けとみなした.

まず, 初心者向け料理レシピのテキスト表現には以下の特徴が見られた.

- 単一の素材に関する調理動作において, 素材そのもの及び調理動作について詳細に説明
- 複数の素材に関する調理動作において, 動作前後の素材の状態について詳細に説明

次に, 表 1 に示すように, 初心者向け料理レシピの画像・映像表現には, 調理動作の種類に応じて異なる使われ方が見られた. ここで調理動作の種類は, 浜田

(注1): 任天堂(株), “しゃべる! DS お料理ナビ”.

(注2): ECCBR2008, ICCBR2009, 2010 国際会議に併設されて開催.

(注3): (株)エルネット, “ボブとアンジー,” <http://www.bob-an.com/>.

らによる分類 [1] を準用した．特に，標本数が比較的多い調理動作には，以下の特徴が見られた．

- 画像：主に「加熱」動作を補足
- 連続画像^(注4)：主に「切削」動作を補足
- 映像：主に「混合」動作を補足

3. 料理レシピを分かりやすくするための補足方法

3.1 分かりにくい表現の検出

2. の分析結果から，以下の条件を満たす語の対を「分かりにくい表現」として検出する．ここでは，大量の料理レシピ中で出現頻度が低く，なじみがない調理動作や素材は初心者にとって分かりにくいと仮定した．

(1) 我々が既に提案した手法 [8] により，「調理動作 (動詞)」 t_v と「素材 (名詞)」 t_n の対を作る．ここで，形態素解析には「MeCab」^(注5)を用いた．

(2) 語 t_v, t_n の一般的な出現頻度 f_v, f_n に基づき，次式で定義される語の対の顕著度を計算する．

$$R(t_v, t_n) = 1 / (w_v f_v + w_n f_n) \quad (1)$$

ここで w_v, w_n は， t_v の調理動作の種類に応じて，以下のように経験的に設定した．

- t_v が単一素材に対する調理動作 (「調理動作 (単一)」): $(w_v, w_n) = (0.1, 0.9)$
- t_v が複数素材に対する調理動作 (「調理動作 (複数)」): $(w_v, w_n) = (0.3, 0.7)$

(3) $R(t_v, t_n)$ がしきい値 θ_R より高い場合に， (t_v, t_n) を分かりにくい表現と判断する．

ここで，「調理動作 (単一)」，「調理動作 (複数)」，「素材」の分類及び一般的な出現頻度は，以下に示す方法で事前に作成した辞書により求める．

- 「調理動作 (単一・複数)」: 浜田らが作成した辞書 [1] に基づいて分類し，大量の料理レシピの調理手順における出現頻度を計数

- 「素材」: 大量の料理レシピの素材一覧における出現頻度を計数

オンライン料理レシピサイト^(注6)から収集した料理レシピ 6,779 件から，2,202 種類の素材 (名詞)，1,568 種類の調理動作 (動詞；単一 795 種類，複数 773 種類)，197 種類の副詞，118 種類の形容詞，7,574 種類のその他の名詞の出現頻度を得た．

3.2 マルチメディア補足候補情報の準備

分かりにくい表現を補足するためには，補足候補情報を事前に準備する．以下，テキストと画像・映像の場合に分けて説明する．なお，語の分類及び使用した

辞書は，3.1 で紹介したものと同じである．

3.2.1 テキストによる補足候補情報

ある素材 t_n が多数の料理レシピにおいて同様の調理動作 t_v を施されるとき，共通して現れる修飾語 t_{m_i} が語の対 (t_n, t_v) に対する典型的な説明であると考え，補足候補情報として抽出する．

(1) 入力された料理レシピ中の t_n について，類似した調理過程を経るレシピ r_i を大量の料理レシピから抽出する．ここで，調理過程とは料理レシピ r_i において t_n に施される調理動作の系列 $P(t_n, r_i) = \{t_{v_1}, t_{v_2}, \dots\}$ とする．また，2 件の料理レシピ r_1, r_2 間の t_n に対する調理過程の類似度は，次式のように正規化編集距離で定義する．

$$d(P(t_n, r_1), P(t_n, r_2)) = \frac{d_e(P(t_n, r_1), P(t_n, r_2))}{\max(\|P(t_n, r_1)\|, \|P(t_n, r_2)\|)} \quad (2)$$

ここで， $d_e(\cdot, \cdot)$ は，2 系列間の DP マッチングによる編集距離を示す．

得られた距離がしきい値 θ_d よりも短いとき，2 件の料理レシピにおいて t_n に対する調理過程が類似していると判断する．また，このようにして得られた類似した料理レシピの集合を t_n の CCORS (Common Cooking Operation Recipe Set) と呼ぶ．

(2) t_n の CCORS 中の全ての (t_n, t_v) について， t_n, t_v の間に現れる全ての修飾語 (副詞が形容詞) t_{m_i} を抽出する．これは，日本語の一般的な主述構造による．様々な例外による検出漏れもあり得るが，大量のレシピから収集するため，無視できると考える．

(3) 全ての t_{m_i} について，次式で定義される CCORS における共通度を計算する．

$$C(t_{m_i}; t_n, t_v) = \frac{t_n, t_{m_i}, t_v \text{ の共起頻度}}{t_n, t_v \text{ の共起頻度}} \quad (3)$$

(4) 共通度が θ_C 以上の修飾語を抽出し，補足候補情報を含めた語の組 (t_n, t_{m_i}, t_v) を出力する．ただし，分量・時間・温度に関する定量的な情報は状況に強く依存するため，除外する．

3.2.2 画像・映像による補足候補情報

調理動作を分かりやすく説明するためには，テキス

(注4): 手順を説明するために複数のイラスト画像を列挙したものを．

(注5): 京都大学，“日本語形態素解析器 MeCab”

<http://mecab.sourceforge.net/>

(注6): 味の素 (株)，“レシピ大百科”

<http://www.ajinomoto.co.jp/recipe/>

トによる補足に加えて、画像・映像による補足も重要である。本研究では、そのために必要な索引付きの画像・映像は、文献 [11], [12] などの方法によってテレビジョン放送やインターネット上の素材をもとに既に存在すると仮定する。ただし、以下の実験においては、人手で収集・索引付けした画像・映像を用いた。

3.3 マルチメディア情報による補足

まず、入力された料理レシピ中の表現 (対 (t_n, t_v)) に対して補足候補情報の組 (t_n, t_m, t_v) が存在する場合に t_m をテキスト中に補足する。

次に、表現 (対 (t_n, t_v)) に対して、適切な画像・映像が存在する場合に、以下の規則に従って、 t_v の種類に応じて画像・連続画像・映像のいずれかを補足する。

- 画像：「加熱」動作を補足
- 連続画像：「切削」動作を補足
- 映像：「混合」動作を補足

4. 実験

4.1 分かりにくい表現の検出

4.1.1 実験条件

被験者実験により、分かりにくい表現の検出手法を評価する。対象とする料理レシピは、ウェブサイト^(注3)から 11 件の上級者向けのもの（一般向けレシピのうち特に難易度が高いもの）を選んだ。被験者として初心者（日常的に料理をしない者）12 名を集め、11 件の料理レシピ全てを見せ、実際に調理することを考えながら、分かりにくい表現を指摘させた。なお、表現の顕著度に関するしきい値 θ_R は、2. において分析した子供向けのものとは大人向けのものにおける表現の違いに関する顕著度の分布（図 1）を参考に、これらを最もよく検出できる値に設定した。

4.1.2 結果

提案手法は、分かりにくい表現を 62 箇所検出した。被験者ごとの再現率、適合率は各々、55-88%（平均 69%）、11-40%（平均 25%）だった。指摘された表現

を分析すると、被験者ごとに大きく異なっていた。このことから、分かりにくい表現をより良く検出するためには、利用者ごとの知識や経験を考慮する必要があることが示唆された。

4.2 テキストによる補足候補情報の抽出

4.2.1 実験条件

3.1 において辞書を作成する際に使用した 6,779 件のレシピから選んだ 3 種類の CCORS（肉じゃが 14 件、コロッケ 26 件、ケーキ 80 件）から抽出された補足候補情報について、人手で分析をした。なお、調理過程の類似度に関するしきい値 $\theta_d = 0.8$ 、修飾語の共通度に関するしきい値 $\theta_C = 0.5$ は経験的に設定した。

4.2.2 結果

合計 185 件の補足候補情報が抽出された。我々はそのうち 136 件（73.5%）について、適切と判断した。表 2 に適切と判断されたものの一部を例示する。

結果を分析すると、しきい値 θ_d により制御される CCORS の大きさが大きく影響を及ぼすことが分かった。つまり、CCORS が大きいと雑音が増えるため、補足候補情報の質が低下し、逆に小さいと補足候補情報がほとんど得られない。このことから、より良い補足候補情報を抽出するためには、素材ごとに θ_d を変化させる必要があることが分かった。

4.3 マルチメディア情報の補足

4.3.1 実験条件

最後に、提案手法によりテキスト料理レシピ 4 件（肉じゃが、カントリーアップルケーキ、タイのけんちん蒸し、エビとソラマメのペニエ）について実際にテキスト・画像・映像からなるマルチメディア情報を補足し、被験者実験により評価した。図 2 に補足例を示す。被験者として初心者（日常的に料理をしない者）8 名を集め、4 件の料理レシピ全てを見せ、実際に調理することを考えながら、補足により分かりやすくなっ

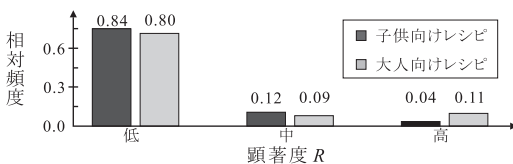


図 1 子供向けレシピと大人向けレシピに出現した 1,184 件の表現 (t_v, t_n) に関する顕著度 R の頻度分布

Fig.1 Frequency distribution of rarity R for 1,184 expressions (t_v, t_n) that appeared in the recipes for children vs. grown-ups.

表 2 抽出された補足候補情報の例

Table 2 Example of the extracted supplemental information candidates.

レシピ	素材 (t_n)	調理動作 (t_v)	補足表現 (t_m)
肉じゃが	ニンジン	むく	皮を
	ジャガイモ	切る	一口大に
	ジャガイモ	むく	皮を
コロッケ	タマゴ	割り混ぜる	ボウルの中で
	キャベツ	ちぎる	手で
	タマネギ	炒める	フライパンで
ケーキ	小麦粉	混ぜる	ボウルの中で
	バター	焼く	オープン

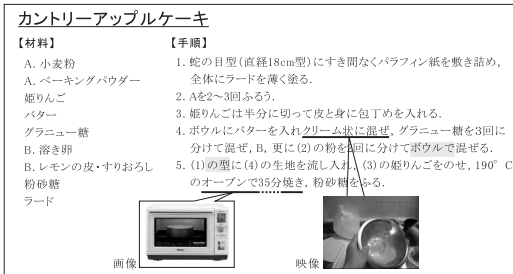


図 2 提案手法により補足された料理レシピの例。網掛けされたテキスト、吹き出し中の画像、映像が補足された。テキストレシピの出典：味の素（株），“レシピ大百科”^(注6)

Fig. 2 Example of a cooking recipe supplemented by the proposed method. The highlighted texts, the image and the video clip in the balloons were supplemented.

表 3 補足された料理レシピの評価：被験者の回答数
Table 3 Evaluation of the supplemented cooking recipes: Number of responses from the subjects.

レシピ	肉じゃが	ケーキ	けんちん	ベニエ
改善された	8	6	8	7
改悪された	0	0	0	1
どちらとも言えない	0	2	0	0

たかを判定させた。

また、テキストの補足の場合は、意味が変わらない範囲で、周囲の語句と整合するように、人手で表現を調整した。

4.3.2 結果

表 3 に被験者による評価結果を示す。ほとんどの被験者が提案手法による補足を「改善された」と判断したが、「どちらともいえない」、「改悪された」と判断することもあった。その理由として、特に調理動作を行う容器が補足された場合にかえて冗長で読みにくくなる点と、依然として一部の表現が分かりにくい点が指摘された。この結果からも、利用者ごとの知識や経験を踏まえた補足が必要なが改めて示唆された。

5. むすび

本論文では、既存のテキスト料理レシピ中の初心者に分かりにくい表現を検出し、テキスト・画像・映像のマルチメディア情報により補足する方法を検討した。実験の結果、提案手法により料理レシピの記述をおおむね分かりやすくできる可能性が示唆された。

今後の課題としては、利用者ごとの知識や経験を踏まえた検出・補足手法を考える必要がある。

謝辞 本研究の一部は、文部科学省科学研究費補助

金（21013022）による。実験の被験者各位及び議論に加わって下さった小幡佳奈子氏に感謝する。

文 献

- [1] 浜田玲子, 井手一郎, 坂井修一, 田中英彦, “料理テキスト教材における調理手順の構造化,” 信学論 (D-II), vol.J85-D-II, no.1, pp.79–89, Jan. 2002.
- [2] 三浦宏一, 高野 求, 浜田玲子, 井手一郎, 坂井修一, 田中英彦, “料理映像の構造解析による調理手順との対応付け,” 信学論 (D-II), vol.J86-D-II, no.11, pp.1647–1656, Nov. 2003.
- [3] R. Hamada, J. Okabe, I. Ide, S. Satoh, S. Sakai, and H. Tanaka, “Cooking Navi: Assistant for daily cooking in kitchen,” Proc. 13th ACM Int. Multimedia Conf., pp.371–374, Nov. 2005.
- [4] 志土地由香, 出口大輔, 高橋友和, 井手一郎, 中村裕一, 村瀬 洋, “料理レシピ中の初心者に理解困難な表現の抽出,” 信学技報, MVE2009-70, Nov. 2009.
- [5] 志土地由香, 出口大輔, 高橋友和, 井手一郎, 中村裕一, 村瀬 洋, “料理レシピをわかりやすくするための理解困難な表現の補足,” 信学技報, MVE2009-145, March 2010.
- [6] I. Ide, Y. Shidochi, Y. Nakamura, D. Deguchi, T. Takahashi, and H. Murase, “Multimedia supplementation to a cooking recipe text for facilitating its understanding to inexperienced users,” Proc. 12th IEEE Int. Symposium on Multimedia, pp.242–247, Dec. 2010.
- [7] K. Miyawaki and M. Sano, “A cooking support system for people with higher brain dysfunction,” Proc. ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities, pp.47–52, Oct. 2009.
- [8] Y. Shidochi, T. Takahashi, I. Ide, and H. Murase, “Finding replaceable materials in cooking recipe texts considering characteristic cooking actions,” Proc. ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities, pp.9–14, Oct. 2009.
- [9] A. Blansch e, J. Cojan, V. Dufour-Lussier, J. Lieber, P. Molli, E. Nauer, H. Skaf-Molli, and Y. Toussaint, “Taaable 3: Adaptation of ingredient quantities and of textual preparations,” 18th Int. Conf. on Case-Based Reasoning Workshop Procs., pp.189–198, July 2010.
- [10] M. Minor, R. Bergmann, S. G org, and K. Walter, “Adaptation of cooking instructions following the workflow paradigm,” 18th Int. Conf. on Case-Based Reasoning Workshop Procs., pp.199–208, July 2010.
- [11] R. Hamada, S. Satoh, and S. Sakai, “Detection of important segments in cooking videos,” Proc. IEEE Workshop on Content-based Access of Image and Video Libraries 2001, pp.118–123, Dec. 2001.
- [12] 柴田知秀, 加藤紀雄, 黒橋禎夫, “言語情報と映像情報の統合による物体のモデル学習と認識,” 情処学論, vol.49, no.3, pp.1451–1464, March 2008.

(平成 22 年 9 月 23 日受付, 23 年 2 月 2 日再受付)