

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 5 日現在

機関番号：14301

研究種目：若手研究(B)

研究期間：2011～2014

課題番号：23700144

研究課題名(和文)調理をしながら調理法を説明した音声からのレシピテキストの自動生成

研究課題名(英文)Recipe dictation system from verbal illustration of how to cook while cooking

研究代表者

山肩 洋子 (Yamakata, Yoko)

京都大学・情報学研究科・特定准教授

研究者番号：60423018

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究では、調理者が調理をしながらその調理法を説明した音声から、レシピテキストを自動生成する手法の研究を行った。レシピテキストにおいては、手順番号の手順で生じた中間食材を、「**1**」に「**2**」を混ぜます」というように手順番号で照応するのが一般的である。しかし音声でそのような中間食材を表現するときは、『さっき切った野菜』というように、適当な呼称を用いることが多い。そこで、調理観測映像から得た情報と調理者との対話から得た情報により、食材の調理状況を認識することで、調理者が食材を音声で表現した際に用いた呼称を照応表現に自動変換する手法を開発し、学会発表を行った。

研究成果の概要(英文)：This research constructed a method for generating a cooking procedural text automatically by recognizing a cooker's verbal explanation of how to cook. Usually, on a Japanese recipe, an intermediate food that is manipulated on the process (1), for example, is designated by the number of the process such as "mix (1) to (2)." However, people never call an intermediate food by such a process number but call it as "the vegetable I cut" or "the ingredients of a NIKUJAGA." Therefore, this research discovered a human wording rule for designating an intermediate object by analyzing cooking videos and interviews with many cooks, and constructed a method to translate automatically a cooker's designation in speech to suitable wording in text.

研究分野：メディア情報処理

キーワード：メディア情報学 ヒューマンインターフェイス 音声対話システム 映像認識 自然言語処理

1. 研究開始当初の背景

近年、健康などの観点から、食事記録への注目が集まっている。栄養学的には、料理名だけでなく、どの材料がどれだけ使われ、どう調理されたかを記録することが重要である。そこで、家庭の調理者が調理しながら調理法を説明した音声から、レシピを自動生成することを考えた。

しかし、調理者の音声をテキストに書き起こすだけでレシピテキストになるわけではない。レシピテキストでは、例えば『皮をむいて角切りにしたリンゴ』のような中間食材は、その食材が手順番号 の手順の結果生じた食材であれば、「 」と番号により照応する。これはテキストは が何かを遡って確認できる一覽性を持つためである。一方、実世界でその食材を加工している調理者は、その中間食材がどの材料からどのような加工を経てきたかといった履歴に基づきその中間食材を認識しているため、「リンゴ」や「先ほど切った物」など、その中間食材の調理履歴に基づく呼称で参照する。このように、中間食材に対する認識の機構は、言語情報から認識する場合と、実世界における視覚情報から認識する場合とで異なると考えられ、読者の理解の仕方に沿ったレシピテキストを生成するためには、調理者が与えた呼称を手順番号による照応表現に変換する必要があると考えた。

2. 研究の目的

本研究では、調理のような人間の創作活動において、一時的にしか存在しないため固有の名前を持たない中間物に対し、実世界でそれを加工している人間がその物体を指示する際に使う音声表現を、文章を読んでその状況を理解する人間にとって理解しやすい言語表現に変換する手法の研究を通して、創作活動で扱う物体に対する人間の「実世界における視覚情報からの認識の機構」と「言語情報からの認識の機構」の相違と関係性を明らかにすることを目的とする。

3. 研究の方法

本課題では、調理者との対話と調理観測映像から調理履歴を獲得し、中間物の呼称を手順番号による照応表現に変換して、レシピテキストを自動生成する機構について研究した。カメラによる調理観測映像を処理し、また調理者の音声から得た呼称を解釈して、両者を調理履歴と対応付ける部分は従来研究で取り組んだ。本課題ではこれらの成果を基盤とし、次の3つの課題を扱った。

(1) 課題1：調理状況を把握するための質問生成

Web 上にある大量のレシピテキストを解析することで、加工の種類ごとに典型的な説明文のテンプレートを導出し、そこから調理状況を把握するための質問項目を生成する。たとえば加熱加工が「[食材]を[状態]になるま

で加熱します。」と説明されることが多ければ、[食材]や[状態]といったスロットが埋まるよう質問を生成する。また、レシピテキスト解析結果と調理履歴から、このスロットを埋めるべき食材の呼称や状態名の候補を導出し、調理者の応答発話認識のための言語モデルや、その加工動作を画像認識する際の注目特徴の決定に利用する。

(2) 課題2：対話による調理履歴の生成

(1)で生成された質問に対する調理者の応答から食材の呼称を抽出し、これを解釈することで、その食材の過去や未来における名前を特定する。それと同時に、調理観測映像中で各食材を追跡することで、各時刻において調理者が扱っている食材が過去および未来におけるどの物体と同一かを判別して、材料から料理の完成品に至る調理履歴を生成する(この調理履歴は食材ノードと加工ノードからなるフローグラフで記述される)。ここで従来研究では、調理者が呼んだ「リンゴ」がある物体(例えば『角切りにしたリンゴ』)のことと同定できるのは、物体追跡によりその『角切りにしたリンゴ』が「リンゴ」と呼ばれる材料と時間的に連続していることを知っているためであった。本研究ではこれとは逆に、ある物体が「角切りにしたリンゴ」と呼ばれたら、それは未加工の「リンゴ」と同一物体と考えるというように、呼称解釈に基づき二つの物体の時間的な連続性を獲得する手法を提案する。

(3) 課題3：照応表現生成

(2)の手続きにより、(1)における[食材]のスロットが、調理履歴中の一つの食材(食材ノード)と対応付けられる。ここで、この食材が未加工の材料でなかった場合は、調理者が呼んだ「角切りにしたリンゴ」といった呼称を、その食材を生成した加工の手順番号に置き換えることで照応表現に変換する。これによりスロットを埋めることで、典型的な調理説明文を生成する。

4. 研究成果

(1) 課題1：調理状況を把握するための質問生成

レシピテキストの自動生成には、テキストの構造を解析したり、高精度な自動音声認識を構築する必要があるため、レシピテキストに適応した辞書や言語モデルが有用である。そこでレシピテキストから半自動により固有表現を抽出することにより、食材名や加工動作名などの固有表現辞書と言語モデルの構築を行った。具体的には、まずレシピテキスト中の各単語が、食材名や加工動作名などの種類の固有表現であるかを推定するため、Web より大量のレシピテキストを収集し、一部を手動でラベリングして京都テキスト解析ツールキット(KyTea)によりモデルを学習した。この結果、未知のレシピについても8割程度の精度で固有表現の種類を推定できることが分かった。次に、レシピテキストを

構文解析して、加工を示す動作にかかる目的語や形容表現を特定した。これを同種の加工を示す動詞間でまとめることで、その種類の加工の説明文のテンプレートを導出した。この用語リストには加工の種類に応じた生起確率を付与しておくことで、調理者の発話認識における言語モデルとして利用できるだけでなく、音声対話による質問を生成するためのテンプレートとしても利用できる。なお本課題は、自然言語処理部分について京都大学学術情報メディアセンター准教授 森信介先生ご自身の研究内容と重なる部分が多かったことから、研究協力体制を築き研究を効率的に進めることができた。

(2) 課題2：対話による調理履歴の生成

(1)で構築した言語モデルにより、調理者が行っているレシピの手順を自動的に構造化すると同時に、調理者の発話から調理進行のどの部分かを推定する機構を構築した。具体的には、まず(1)で構築した言語モデルを用いて、レシピの各手順説明文を係り受け解析し、各手順間を連結することにより、材料が葉ノード、各加工が中間ノード、料理の完成品が根ノードとなるようなフローグラフに変換した。次に、レシピテキストの解析結果と実際の動作の齟齬を解消する仕組みを導入した。たとえば、レシピに「 を炒めます。さらに を混ぜ合わせます。」と書かれている場合、「混ぜ合わせる」作業を行っているときは、その前の「炒める」作業が継続している。また、下ごしらえにおいては省略が頻発する。たとえば「じゃが芋の皮を剥く」のは当たり前であり、よってこのような説明は省かれる場合が多いが、実際には存在する。そこで、レシピから生成したルールを用いることで、これらの齟齬を可能な限り解消した。最後に、調理中に発話された説明発話に対して、前述の言語モデルにより解析し、該当する動作をフローグラフ上で探索することで、現在行っている手順を特定するアルゴリズムを構築した。

(3) 課題3：照応表現生成

まず、レシピの手順説明文を作業フローグラフの形式に自動変換する手法の精度向上を試みた。フローグラフのフォーマットを整理し、手動でフローグラフに変換したレシピを200用意してモデル学習することにより、6割の精度が達成できることが分かった。次に、得られたフローグラフとレシピテキストの関係を分析することで、照応表現を含む文章を生成するルールを導出した。これにより、「リンゴとミカンには皮をむいて一口大に切ります」というように、異なる種類の食材でも同じ加工が加えられる場合には一文にまとめたり、「 を に混ぜてください」というように、同じ作業の対象となっている食材を照応表現に変換して文を生成するシステムを構築した。本研究は平成26年度まで続く予定であった。しかしながら、真に実用可能な程度の高精度

なフローグラフ化を実現するためにはさらなるコーパスの増量や手法の改善が必須であり、その経費を確保したいこと、また、ここ数年のレシピ関連業界の盛り上がりを受けて、社会に向けて提案手法を実用面でアピールできる研究を行いたいと考えたことから、最終年度を待たず今年度での基盤研究(B)の提案を決意し、採択された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

Hidenori Tsuji, Yoko Yamakata, Takuya Funatomi, Hiromi Hiramatsu, Shinsuke Mori, "A Platform for Facilitating Video-based Cooking Communication by Multimedia Technologies," Journal of Multimedia Processing and Technologies, Vol. 3, Issue 2, pp. 81-101, 2013.

山肩洋子, "情報分野における食メディア研究," 月刊 画像ラボ, Vol. 24, No.7, pp.1-8, 2013.

山肩洋子, 角所考, 美濃導彦, "呼称により指示された中間食材を同定するための画像認識モデルおよび呼称解釈法の提案," 電子情報通信学会論文誌, Vol. J94-A, No.7, pp. 519-531, 2011.

橋本敦史, 森直幸, 舩富卓哉, 山肩洋子, 椋木雅之, 角所考, 美濃導彦, "把持の順序と外見の変化モデルを利用した調理作業における食材追跡," 電子情報通信学会論文誌, Vol. J94-A, No.7, pp.509-518, 2011.

[学会発表](計12件)

Yoko Yamakata, Shinji Imahori, Yuichi Sugiyama, Shinsuke Mori, and Katsumi Tanaka, "Feature Extraction and Summarization of Recipes Using Flow Graph," SocInfo 2013, LNCS 8238, pp. 241-254, 2013. 査読有り (Full paper 採択率 22%).

森信介, 山肩洋子, 笹田鉄郎, 前田浩邦, "レシピテキストのためのフローグラフの定義," 自然言語処理研究発表会, Vol. 2013-NL-214, No. 13, pp.1-7, 2013.

山肩洋子, 今堀慎治, 杉山祐一, 田中克己, "レシピフローグラフを介したレシピ集合の要約と特徴抽出," 電子情報通信学会 技術研究報告, Vol. 113, No. 214, DE2013-40, pp. 43-49, 2013.

前田浩邦, 山肩洋子, 森信介, "レシピテキストからのフローグラフコーパス作成," 電子情報通信学会 技術研究報告, Vol. 113, No. 214, DE2013-39, pp. 37-42, 2013.

杉山祐一, 山肩洋子, 田中克己, “手順情報としてのレシピデータに対する類似レシピの要約と微小で重要な差異の発見”, 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), No. D3-5, 2013.

Shinsuke Mori, Tetsuro Sasada, Yoko Yamakata, Koichiro Yoshino, “A Machine Learning Approach to Recipe Text Processing,” Cooking with Computers Workshop, pp.1-6, 2012. 査読有り

Hidenori Tsuji, Yoko Yamakata, Takuya Funatomi, Hiromi Hiramatsu, Shinsuke Mori, “IwaCam: a Multimedia Processing Platform for Supporting Video-Based Cooking Communication,” International Conference on Future Generation Communication Technology, pp.109-116, 2012. 査読有り

山肩洋子, 橋本敦史, 船富卓哉, 平松紘実, 辻 秀典, 森信介, “IwaCamを用いたメディア処理による遠隔コミュニケーションの支援 同時調理を題材とした映像と音声の役割に関する考察,” 電子情報通信学会データ工学研究会, Vol. 112, No. 75, pp. 7-12, 2012. 井上仁, 中村和晃, 船富卓哉, 橋本敦史, 上田真由美, 山肩洋子, 美濃導彦, “画像・振動音・荷重データを統合的に用いた食材認識,” 電子情報通信学会マルチメディア・仮想環境基礎(MVE)研究会, Vol. 111, No. 478, pp. 37-42, 2012.

Atsushi Hashimoto, Jin Inoue, Kazuaki Nakamura, Takuya Funatomi, Mayumi Ueda, Yoko Yamakata, Michihiko Minoh, “Recognizing Ingredients at Cutting Process by Integrating Multimodal Features,” Proceedings of the ACM Multimedia 2012 Workshop on Multimedia for Cooking and Eating Activities, pp.12-18, 2012. (CEA '12 Best Paper Award) 査読有り

Yoko Yamakata, Yoshiki Tsuchimoto, Atsushi Hashimoto, Takuya Funatomi, Mayumi Ueda, Michihiko Minoh, “Cooking Ingredient Recognition based on the Load on a Chopping Board during Cutting,” 3rd Workshop on Multimedia for Cooking and Eating Activities (CEA2011), in conjunction with the IEEE International Symposium on Multimedia, pp. 381-386, 2011. 査読有り

土本良樹, 橋本敦史, 船富卓哉, 山肩洋子, 上田真由美, 美濃導彦, “調理における切断加工時の荷重特徴を用いた食材認識,” 電子情報通信学会マルチメディア・仮想環境基礎(MVE)研究会,

Vol. 110, No. 456, pp. 55-60, 2011.

〔図書〕(計1件)

山肩洋子, “料理と食を取り巻く情報メディア技術,” 電子情報通信学会 技術研究報告(招待講演・総説), Vol. 112, No. 225, PRMU2012-66, pp. 79-86, 2012.

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

形態素解析器 KyTea (京都テキスト解析ツールキット)のレシピへの適用モデルを提供

<http://www.phontron.com/kytea/index-ja.html>

レシピの手順文書からレシピ特有の固有表現を抽出するツールの提供

<http://plata.ar.media.kyoto-u.ac.jp/mori/research/topics/NER/>

レシピの手順文書を作業フローグラフに手作業で書き換えたフローグラフコーパスを提供

<http://plata.ar.media.kyoto-u.ac.jp/mori/research/NLR/FGC/main.html>

6. 研究組織

(1) 研究代表者

山肩 洋子 (Yoko Yamakata) 京都大学・大学院情報学研究科・特定准教授

研究者番号: 60423018