

機関番号：14301

研究種目：若手研究(B)

研究期間：2008～2010

課題番号：20700120

研究課題名(和文)

音声対話における調理法教示のための中間食材の認識および呼称の決定・解釈

研究課題名(英文)

The study of wording and word interpretation for intermediate object on cooking instructional system with spoken dialogue

研究代表者

山肩 洋子 (YOKO YAMAKATA)

京都大学・大学院情報学研究科・准教授

研究者番号：60423018

研究成果の概要(和文)：

一般の調理者の多くは、新しい料理に挑戦する際、料理本などのレシピテキストを読むのが一般的であるが、テキストを読みながら包丁や火を扱うのは危険なため、レシピテキストに基づき調理法を音声で教示するシステムの構築を目指した。これを実現するには、レシピテキストをただ音声で読み上げればいわけではなく、レシピテキストにおいて「1を2に混ぜてください」というように手順番号で示されている中間食材に対して、「先ほど切った玉葱」や「玉葱と人参の混ぜたもの」というように、どの中間食材を指しているか調理者が容易に判別できる呼称を与えなければならない。そこで本研究では、調理者が中間食材をどのようなものと認識しているかを理解し、また調理者との間で共通の名前空間を確立することで、中間食材の呼称を決定する機構の研究を行った。

この研究の成果により、中間食材は『直前の加工名』『構成材料名』『料理名』のほか、『器』『位置』『よく使われる中間物名(生地・タネ・出汁等)』などが組合さった呼び方がなされ、よってキッチン状況をリアルタイムに認識することで呼称の自由度が向上することが示された。また、本研究で提案したモデルに従い言語解釈することにより、調理中に現れたすべての食材のうち93%の食材について正しく同定できたことを示した。

さらに、料理コミュニケーション支援ソフトウェアIwacamの開発に参加し、Iwacamに音声対話システムとこれまで開発してきた食材認識技術を組み込んだ。これにより、カメラやマイクにより調理を観測し、その情報を逐次解析したり、音声対話により認識誤りを補正することで、各時刻に調理台上にどのような食材が存在するか、どのような加工を加えられているか、それは過去のどの物体と同一のものかを判別するシステムを実装した。

研究成果の概要(英文)：

Most of home chef refers a cooking recipe text when he/she try to cook a new menu. However cooking while reading a recipe is sometimes very danger because he/she has to handle a knife and fire in the cooking. Therefore, we aim to construct a recipe guidance system that instructs how to cook according to its recipe text. To realize such recipe guidance system, it is no enough to read the recipe book. For example, it is very common that a recipe gives such instruction as “mix (1) to (2).” This sentence means an intermediate object given by process (1) would be mixed with an intermediate object given by process (2). If the system just reads this sentence with a voice, a chef might not decide which objects the system designates because he/she does not know/remember which object is given by process (1)/(2). In this case, the system should give reasonable name such as “cut onion by the previous process” and “a mix of onion and carrot”. Such reasonable name should be according to like what the chef understand the object. Therefore, we propose a method to give reasonable name according to the model of user understanding toward an intermediate food object in a cooking process.

In this research, we presented such intermediate objects is tend to be designated by

combination of “the name of the last cooking action toward the object”, “the name of its components”, “the name of meal the object going to be”, “the name of vessel”, “the name of position”, and “the name of milestone (ex. “sauce”, “dough). We proposed a method to specify the designated intermediate object by a model of human naming toward intermediate object. The proposed model archives 93% accuracy of object specification.

According to the model of human naming toward intermediate object, we construct a recipe guidance system based on IwaCam, the a live video cooking communication software. IwaCam aims to enable cooperative cooking between distant kitchen spaces.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,300,000円	390,000円	1,690,000円
2009年度	1,000,000円	300,000円	1,300,000円
2010年度	900,000円	270,000円	1,170,000円
年度			
年度			
総計	3,200,000円	960,000円	4,160,000円

研究分野：総合領域

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：ヒューマンインターフェイス

1. 研究開始当初の背景

一般の調理者の多くは、新しい料理に挑戦する際、料理本などのレシピテキストを読むのが一般的であるが、テキストを読みながら包丁や火を扱うのは危険なため、レシピテキストに基づき調理法を音声で教示するシステムの構築を目指した。これを実現するには、レシピテキストをただ音声で読み上げればいわけではなく、レシピテキストにおいて「1を2に混ぜてください」というように手順番号で示されている中間食材に対して、「先ほど切った玉葱」や「玉葱と人参の混ぜたもの」というように、どの中間食材を指しているか調理者が容易に判別できる呼称を与えなければならない。そこで本研究では、調理者が中間食材をどのようなものと認識しているかを理解し、また調理者との間で共通の名前空間を確立することで、中間食材の呼称を決定する機構の研究を行った。

2. 研究の目的

1. 時間的・空間的コンテキストに基づく中間物認識モデル 「今指示しようとしている中間物は、過去のどの物体と同一か、また未来のどの物体に成ると推定されるか」という、物体の時間的連続性によって結び付けられる過去・未来の情報（時間的コンテキスト）と、現時点における物体の色や形などの観測特徴、キッチン上での位置および他の物体に対する相対位置などの情報（空間的コンテキスト）により統合的に中間物をモデル化して

認識する手法を提案する。

2. 状況に適応的な言語理解モデル 同じ物体を複数の呼称で呼び得るとき、話者がその物体を指示するために選ぶ呼称は、その時の状況において指示対象物体だけがそう呼べ、他の物体はそう呼べないような呼称であると考えられる。そこで本研究では、被験者実験による対話分析に基づき、ロボットが中間物を指示する呼称を、その時点で存在するその他の物体の呼称も考慮して決定する言語モデルを構築する。またユーザが呼んだ呼称はそのような言語モデルにより決められたと考えることで、その呼称で指示された中間物を同定する言語解釈モデルを提案する。

3. 共通理解獲得のための対話管理モデル 指示物体の候補が多数存在する場合、聞き手が指示対象を一度で同定できるような呼称が存在するとは限らない。また、ユーザとロボットは時間的・空間的コンテキストを共有するが、それによる中間物の認識モデルは必ずしも一致しない。そこで本研究では、両者が同じ物体を意図していると信じられるまで対話することで、認識や言語理解の不一致を発見・解消し、完全な共通理解を獲得する対話モデルの提案を行う。

3. 研究の方法

調理観測から音声による物体指示・理解に至るまでの処理の流れを下図に示す。上で述べたように、観測映像処理と音声処理は従来研究の成果を利用することとし、本

研究では、Ⅰ) 中間物の認識モデル化、Ⅱ) 言語理解、Ⅲ) 対話管理の3つの処理に焦点をあて研究する。

Ⅰ) 中間物の認識モデル化

●空間的コンテキストによるモデル化

○色や形、大きさ、テクスチャなど、その物体自身が表出している特徴

○入れられている器(「ボウル」「鍋」等)、キッチン上での位置(「右手前」「左のコンロ」等)、他の食材との相対位置(「食材Aの右横」等)など、他の物体との相対的な関係

●中間物の時間的コンテキストによるモデル化

調理の過去・現在・未来に亘る情報に基づき中間物をモデル化する。

○調理開始時から加工対象となる食材を追跡しつづけることにより、現時点のある食材が、過去の別の時点ではどの食材と同一(あるいは構成部分)であるかを対応付ける

○現時点までの物体の時間的コンテキストをレシピと照らし合わせることで、その物体がこの先どのような操作を受け、どのような名前の物体になるかを予測する

●空間的コンテキストと時間的コンテキストの統合

各時刻での各物体の空間的コンテキストを、物体が時間的に連続していることを手がかりに時間軸方向に関連付ける。これを時空間コンテキストと呼ぶ。

○調理で扱うすべての物体に関する、調理の開始から終了に至るまでのすべての情報のモデル化であるため、これをユーザとロボットが持ちうる共通知識の全集合と位置づける

○このうち、対話に用いられることのない部分を枝狩りすることで、人間の創作活動で扱う物体に対する知識管理及び理解の概形を規定し、創作活動記録のための記述を導出する

Ⅱ) 言語理解

●物体の時間的・空間的コンテキストそれぞれに基づく言語理解とその比較

調理の各時点での中間物を被験者に音声で指示させる実験を行う。

○被験者が与えた呼称を、時間的または空間的コンテキストのどちらに基づくかにより分類し分析することにより、各コンテキストより呼称を生成するルールを導出する

○時間的・空間的コンテキストのそれぞれによる呼称がどのように使い分けられたかを分析することで、言語理解における各コンテキストの関係を明らかにする

●時空間コンテキストにもとづく最適呼称の選定・理解

一つの物体は同時に複数の呼称で呼び得るし、それは同時刻に存在する他の物体にお

いても同様であるから、話者は聞き手が指示対象物体を一意に同定できるよう、指示対象のみに該当し、それ以外の物体には該当しないような呼称を選ぶ必要がある

○被験者は自身が付与した呼称で聞き手が対象を一意に同定できると考えたと考え、その呼称からロボットが正しく指示対象を同定できるような言語理解のモデルを構築する

○ロボットの言語理解モデルに基づき付与した呼称を被験者に提示して指示対象を同定させる。いかなる場面でも被験者が正しく対象を同定できるようモデルを修正し、評価する

Ⅲ) 対話管理

●対話による完全な共通理解の獲得

ユーザとロボットが完全な共通理解を獲得していれば、話者は聞き手が指示対象を同定できているかわかるはずである。そこで、対話を行いながら相手の理解を確認することで、互いの誤解を修正し、ユーザとの完全な共通理解を獲得する機構を実現する。

4. 研究成果

(1) 京都大学学術情報メディアセンター美濃研究室において、マルチメディアデバイスにより調理行動を認識して支援を行うシステム“Smart Kitchen Project”の研究開発に協力し、その中で本研究を施行した。

(2) 一般の調理者が中間食材をどう呼ぶか、どう呼べば判別できるかを調べるため、自動応答システムを人間が演じる Wizard of Oz 方式により、料理を教示するシステムと、それを聞きながら調理する調理者との間のやり取りを調査した。これにより、中間食材は『直前の加工名』『構成材料名』『料理名』のほか、『器』『位置』『よく使われる中間物名(生地・タネ・出汁等)』などが組合された呼び方がなされ、よってキッチンの状況をリアルタイムに認識することで呼称の自由度が向上することが示された。

(3) 時間的コンテキストと空間的コンテキストの両方を関連付けた。これにより、調理で扱うすべての食材(中間物を含む)に関する、調理の開始から終了に至るまでの、名付けに必要なすべての情報をモデル化した。

(4) モデルの妥当性を検証するため、モデルに基づき与えた呼称で人間が対象食材を同定できるかを調べた。その結果、調理中に現れたすべての食材のうち93%の食材について正しく同定できたことを示した。

(5) 料理コミュニケーション支援ソフトウ

エア Iwacam の開発に参加し、Iwacam に音声対話システムとこれまで開発してきた食材認識技術を組み込んだ。これにより、カメラやマイクにより調理を観測し、その情報を逐次解析したり、音声対話により認識誤りを補正することで、各時刻に調理台上にどのような食材が存在するか、どのような加工を加えられているか、それは過去のどの物体と同一のものかを判別するシステムを実装した。

(6) 電子情報通信学会 料理メディア研究会の幹事および副委員長を務め、学会誌小特集の企画及び執筆、論文特集号の企画、および国際ワークショップの企画を行った。これにより、関連研究の情報収集を行うとともに、当該研究及びその成果物の広報活動を行った。

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 3 件)

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. 519-131, 2011. 7.

橋本 敦史, 森 直幸, 船富 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. 7.

山肩 洋子, 船富 卓哉, 上田 博唯, 辻 秀典, 美濃 導彦, 中内 靖, 宮脇 健三郎, 中村 裕一, 椎尾 一郎: 小特集生活に役立つメディア処理—料理行動を科学する— 3. 料理を作る「調理」という複雑な知的創作活動を最新の情報科学で支援する, 電子情報通信学会誌, Vol. 93, No. 1, PP. 39-47, 2010. 1.

[学会発表] (計 10 件)

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,” マルチメディア・仮想環境基礎研究会 (MVE), Vol. 111, No. 478, pp. 37-42, 2012. 3.

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 ,

Workshop on Multimedia for Cooking and Eating Activities (CEA2011) in conjunction with The IEEE International Symposium on Multimedia 2011, 2011.12.

山肩 洋子, “鮭屋における調理場観測映像の解析によるサービス支援”, 第 2 回文化とコンピューティング国際会議共催イベント「食文化に関する学際的探求: 作る, 食べる, もてなす」, 2011.10.

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. 3.

YAMAKATA YOKO, KAKUSHO KOH, MINOH MICHIIHIKO: “Object Recognition based on Object’s Identity for Cooking Recognition Task”, 2nd Workshop on Multimedia for Cooking and Eating Activities: CEA2010, in conjunction with The IEEE International Symposium on Multimedia, PP. 278-283, 2010.12.

山肩洋子: “料理コミュニケーション支援ソフトウェアIwacamの一般家庭への導入事例報告”, 情報科学技術フォーラム講演論文集, Vol. 9, No. 3, 2010.8.

辻 秀典, 山肩 洋子, 船富 卓哉, 近藤 一晃, 出口 大輔, 井手 一郎, 宮脇 健三郎, 佐野 睦夫, “料理映像コミュニケーション基盤ソフトウェアIwaCamを用いた遠隔共同料理実験”, 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. 35 PP. 33-34, 2010. 5.

料理メディア研究専門委員会: 双方向映像配信技術により遠隔地間での共食・共同調理を支援するコミュニケーションツール—基盤ソフトウェアの開発報告—, 2009 年度 HCG シンポジウム, K1, 2009. 12.

Yoko Yamakata, Takuya Funatomi, Koh Kakusho, Michihiko Minoh: A Study of Object Naming According to the Manufacturing Processes in a Cooking Activity, ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities (CEA2009), P. 1-8, 2009. 10.

Atsushi Hashimoto, Naoyuki Mori, Takuya Funatomi, Yoko Yamakata, Koh Kakusho, Michihiko Minoh: Smart Kitchen: A User

Centric Cooking Support System,
Information Processing and Management of
Uncertainty in Knowledge-Based Systems,
pp. 848-854, 2008. 6.

〔図書〕 (計 0 件)

〔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

電子情報通信学会ヒューマン・コミュニケー
ション・グループ 食メディア研究会におい
て副委員長を担当

<http://www.ccm.media.kyoto-u.ac.jp/CEA/>

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・准教授

研究者番号：60423018