

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 14 日現在

機関番号：14301

研究種目：基盤研究(A) (一般)

研究期間：2012～2015

課題番号：24240030

研究課題名(和文) 食材, 道具, 動作の認識を連携させた調理行動の認識

研究課題名(英文) Understanding Human Activity in Kitchen via Food, Tool, and Action Recognition

研究代表者

美濃 導彦 (Minoh, Michihiko)

京都大学・学術情報メディアセンター・教授

研究者番号：70166099

交付決定額(研究期間全体)：(直接経費) 35,500,000円

研究成果の概要(和文)：一般の動物とヒトを隔てる「知能」として、本研究では道具の使用, および論理や推論による問題解決能力を利用して行う「知的行動」の認識を目指し, 題材として「キッチンでの調理」を取り上げ, 道具, 食材, 調理動作の認識, および言語的に記述された行動の理解に取り組んだ。観測データから認識できる「表出動作」のうち, 特に物体の把持・解放を中心として, レシピ中に言語的に記述された「概念的行動」の完了認識や次の行動の予測をする手法を構築した。また, テキストから扱われる食材の種類やそれに対する行動, 目標状態などを概念的行動として自動で獲得し, レシピ内でのそれらの依存関係をワークフローとして抽出する技術を構築した。

研究成果の概要(英文)：This study aimed to understand human activity in kitchen in which various tools are used and decision-making is required; the activity represents human intelligence that separates us from other animals. In “visible actions” in observations, we focused on object grabbing and releasing actions, and developed a method to recognize completion of “invisible activities” that are linguistically described in recipe text and to forecast the forthcoming action intended by humans. We also developed several methods to extract “invisible activities” in recipe texts as workflow graph with its word class (i.e. food, tools, food-states and so on).

研究分野：知覚情報処理・知能ロボティクス

 キーワード：映像, 文章, 音声等認識, 情報図書館学, スマートセンサ情報システム, 知能ロボティクス, 行動認識  
 画像, 文章, 音声等認識

## 1. 研究開始当初の背景

一般の動物とヒトを隔てる「知能」として、本研究では道具の使用、および論理や推論による問題解決能力を利用して行う「知的行動」の認識を目指す。この題材として、本課題では道具を使って加工していく「キッチンでの調理」を取り上げ、道具、食材、調理行動の認識・理解に取り組む。

人の行動を観測、理解する技術として、人の動作の観測、分析、認識や、人物や車の追跡、異常検出など、世界中でさまざまな研究が行われている。例えば HASC (Human Activity Sensing Consortium) では、人間行動理解の研究を加速させることを目的とし、装着型センサによる行動認識、理解を目指す研究者、企業が共同でデータセットの構築を行ったり、それをを用いた認識、理解技術を競ったりしている。

HASC を含め、人の行動認識の研究では、対象として「歩く」「座る」等の動作を主として扱っている。これらの動作は認識の題材としては重要であるが、自らの身体を用いた基本的な動作であり、他の動物にも見られる動作であるという点において、本研究が目指す「知的行動」の認識とは本質的に異なる。観測を通じた知的行動の認識、理解は、知能とは何かに迫る重要な研究である。従来までに取り組みされた数少ない事例の中では本課題と同様に調理を題材としたものが存在する。これは「知的行動」の認識ではなく、その環境を構築するもので観測データを収集して公開するのが目的である。しかし、収集された観測データは小規模で、一貫性もなく、十分に整備されていない。これを用いた研究も数少なく「知的行動」の認識まで踏み込めていない。

## 2. 研究の目的

観測データから認識できる(置く、入れる、混ぜる、等)の動作を表出動作、レシピ等において言語で表現される(炒める、等)の行動を概念的行動と呼ぶことにする(図 1)。知的行動の認識、理解においては、

- 道具の使用: 表出動作の認識には、その動作に関わる物体(食材、道具)の認識が重要
- 論理や推論による問題解決: 表出動作と概念的行動の関係をを用いて、認識した表出動作を介して概念的行動を推定可能

である。そこで本研究では、

- 1) 道具、食材の認識、
- 2) 表出動作と概念的行動の関係の獲得、
- 3) 道具、食材の認識と連携した表出動作の認識、

および

- 4) 概念的行動の認識

に取り組む。

また、認識は与えられたコンテキストにおいて利用されるので、

- 5) 認識結果に基づく調理支援システムを構築し、これを通して認識技術の評価、改

善を行う。

## 3. 研究の方法

### 1) 道具、食材の認識

申請者はこれまで、調理を観測した映像中から、調理者が用いた道具や食材を頑健に抽出したり[研究業績 1]、抽出した道具や食材を追跡することで時間的対応関係を構築したり[研究業績 2]する技術を開発してきた。これをベースに、従来の物体認識技術を応用して、道具の認識に取り組む。一方、食材の認識については、見た目に似たものが多く、また個体差が大きいため、加工前の状態をカメラで観測したとしても認識は容易でない。調理では食材に対してさまざまな加工を施していくため、その過程をカメラや荷重センサなどのさまざまなセンサで観測することにより食材についての特徴や加工による変化も取得して食材を認識する。

### 2-a) レシピの分析による概念的行動に関する言語表現の整理

レシピは調理に関する文章記述であり、Cookpad(<http://cookpad.com/>) や楽天レシピ(<http://recipe.rakuten.co.jp/>) などから大量に獲得できる。例えば、同一のメニューであればレシピは似たような手順になるため、自然言語処理でそこから概念的行動の表現を抽出し、その多様性を獲得する(例えば図 3 の a) 野菜を炒める / b) フライパンを熱し、野菜を入れて、炒める / c) フライパンを熱し、油を引いて、野菜を炒める / など)。これにより、概念的行動の語彙を構築する。

### 2-b) 表出動作や概念的行動のタグ付け、データセットの作成

パターン認識技術の立場から、調理者の動作に対して表出動作の語彙を認識が可能な粒度で設定し、レシピ分析の結果抽出された概念的行動の語彙と共に同一の観測データに対してタグを付与していく(図 3)。これにより、表出動作と概念的行動の関係が明確化される。また、多様な概念的行動の表現についても、表出動作を基準として整理する。以上のような観測データに対してタグが付与されたものをデータセットと呼ぶ。

### 3-a) 道具、食材の認識結果とデータセットに基づく表出動作の認識

タグが付与された観測データを対象として表出動作の認識技術を構築する。調理の観測に用いたさまざまなセンサのデータを活用し、表出動作に用いる道具や加工の対象となる食材の認識結果を手がかりとして、従来の行動認識技術を発展させてゆく。

### 3-b) 表出動作認識の適用による大規模データセットの作成

観測データに対するタグの付与は人手と時間がかかる。先述のような表出動作と概念的

行動の関係を導出するためには、大規模なデータセットが必要となる。そのため、表出動作認識技術、および自然言語処理により抽出した概念的行動の語彙を活用して、大規模データセットの構築の効率化を図る。具体的には、表出動作認識によるタグ付けの半自動化、概念的行動の語彙における時間的順序関係を用いたタグの入力予測などを行う。このようにして作成したデータセットは他の研究者にも有用であるので、公開することを念頭にデータセットを整備する。

#### 4) 表出動作と概念的行動の関係を利用した概念的行動の認識

これまでに構築した道具、食材、表出動作の認識技術、および大規模データセットから獲得した表出動作と概念的行動の関係から、概念的行動の認識に取り組む。

### 4. 研究成果

#### 1) 道具、食材の認識

食材は色素が限られており、一般物体認識で用いられる外見的特徴に乏しい、認識の難しい対象である。また、その形状も個体差が大きい。このような対象を認識するために、調理の初期段階で行われる切削作業の際に生じる音や荷重といった内部破壊時の振動情報を併用した識別技術を構築した。また、道具・食材の両方に関して、比較的規模の大きいデータセットを作成し、深層学習による認識を行い、この認識結果やアノテーション情報などを KUSK Object Dataset の中で公開した。

(<http://kusk.mm.media.kyoto-u.ac.jp/kusk-dataset/kusk-object-dataset/>)

#### 2) 表出動作と概念的行動の関係の獲得

Cookpad データセットに含まれるレシピに対してアノテーションを行い、レシピに特化した言語処理技術の開発を行なった。基本技術である字句解析・構文解析だけでなく、文章から作業の構造を抽出する処理までを自動化し、フローグラフコーパスを作成・公開し、言語で表現される概念的行動の整理を行なった。

(<http://www.ar.media.kyoto-u.ac.jp/data/recipe/>)

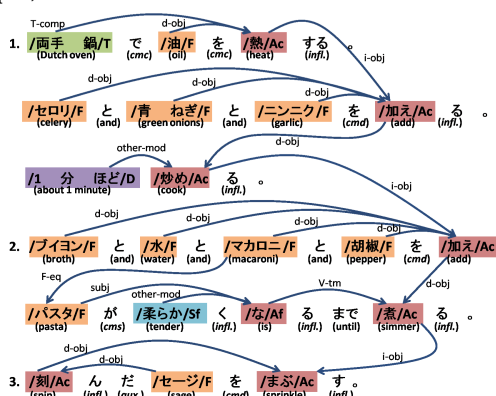


図 1. 抽出されたフローグラフの例

またこのような処理を行なった 20 種類のレシピをそれぞれ複数回調理し、その様子を、マルチセンサ環境を備えたキッチンにより記録し、KUSK Dataset として公開した。  
(<http://kusk.mm.media.kyoto-u.ac.jp/>)

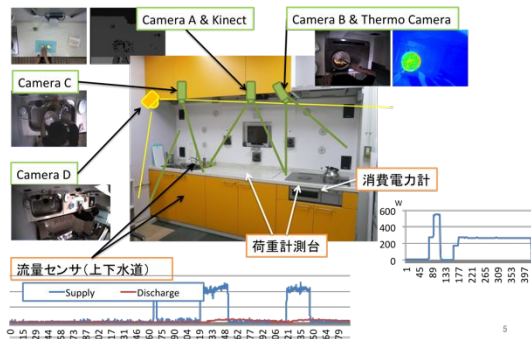


図 2 マルチセンシング環境を備えたキッチン

記録された作業者の行動（表出動作）と文書表現（概念的行動）の対応関係が得られるよう、作業者自らに対応する文を明示させるような実験により、その対応関係も同時に獲得するよう取り計らった。合計で 59 本の調理観測データを得た。また、表出動作のうち、特に物体の把持・開放動作については、半自動での大規模アノテーションを行ない、前述の物体認識結果と合わせて KUSK Object Dataset として公開を行なった。

#### 3) 道具、食材の認識と連携した表出動作の認識

道具・食材と密接に関連する表出動作として、物体の把持・解放に着目し、映像を元に把持・解放を自動で判定する技術、および、荷重センサと組み合わせることでより頑健に把持・解放の検出を行う技術を開発した。

また、複数の食材が合流する「混ぜる」という動作を対象に、混ぜる際に使われる容器の形状に着目した動作認識手法を提案した。

さらに、食材の中でも容器に入れられて少しずつ消費される調味料を対象として、調味料の置き場を秤とし、調味料の出入りや消費量を管理することで、味付け動作の記録を行うシステムを作成した。

この他、把持・解放のタイミングを元に調理映像をセグメントへ分割し、セグメント毎に調理者の動きの特徴量を抽出してクラスタリングすることで、動きに基づいて表出動作をボトムアップに離散化する手法の提案を行なった。セグメントへの分割によって継続時間を考慮した動作クラスタリングを可能とし、従来よりも高度な記号化を実現したが、これらの動作クラスタを概念的行動と対応付けることは今後の課題である。

#### 4) 概念的行動の認識

上述の動作クラスタの利用は間に合わなかったものの、表出動作のうち、特に把持・解

放に着目して、物体の把持・解法履歴に基づいた概念的行動の認識、及び、調理者が次に行おうとしている概念的行動の推定までを行うシステムを開発した。

5) 認識結果に基づく調理支援システム  
対外発表には間に合わなかったが、1~4の成果を統合し、自動でレシピをナビゲーションする支援システムの作成を行なった。このシステムについては実用上の課題が残っており、今後も継続的に開発・機能の改善を続けていく。



図 3. 作成した調理支援システムの画面

#### 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 6 件)

[1] 山崎健史、吉野幸一郎、前田浩邦、笹田鉄郎、橋本敦史、船富卓哉、山肩洋子、森信介、フローグラフからの手順書の生成、情報処理学会論文誌、査読有、57 巻 3 号、2016、849-862、DOI: なし

[2] Kazuaki NAKAMURA, Takuya FUNATOMI, Atsushi HASHIMOTO, Mayumi UEDA, Michihiko MINOH, Development and Evaluation of Near Real-Time Automated System for Measuring Consumption of Seasonings, IEICE TRANSACTIONS on Information and Systems, Vol. E98-D, No. 12, 査読有, 2016, 2229-2241, DOI: 10.1587/transinf.2015EDP7072

[3] 笹田鉄郎、森信介、山肩洋子、前田宏邦、河原達也、レシピ用語の定義とその自動認識のためのタグ付与コーパスの構築、自然言語処理、査読有、22 巻 2 号、2015、104-131、DOI: <http://doi.org/10.5715/jnlp.22.107>

[4] 井上仁、橋本敦史、中村和晃、船富卓哉、山肩洋子、上田真由美、美濃導彦、食材認識のための画像と食材切断時の振動音及び荷重の利用、電子情報通信学会論文誌、査読有、J97-D、No.9、2014、1490-1502、DOI なし

[5] 森信介、自然言語処理における分野適応、人工知能学会誌、査読無、27(4)巻、2012 年、

365-372  
DOI なし

[6] 橋本敦史、船富卓哉、中村和晃、美濃導彦、机上物体検出を対象とした接触理由付けによる誤検出棄却、電子情報通信学会和文論文誌、査読有、J95-D、2012、2113-2123  
DOI なし

[学会発表](計 27 件)

[7] 松村優樹、橋本敦史、森信介、船富卓哉、飯山将晃、美濃導彦、経路探索に基づく映像とワークフローの対応付け、電子情報通信学会マルチメディア仮想環境基礎研究会、2016年3月7日、名桜大学(沖縄県名護市)

[8] 橋本敦史、飯山将晃、森信介、美濃導彦、KUSK Object Dataset: 調理作業中の物体への接触履歴データセットの作成、電子情報通信学会データ工学研究会、2015年9月25日、神奈川工科大学横浜オフィス(神奈川県横浜市西区)

[9] Hirokuni Maeta, Tetsuro Sasada, Shinsuke Mori, A Framework for Procedural Text Understanding, The 14<sup>th</sup> International Conference on Parsing Technologies, July 22, 2015, The Auditorium of the University of the Basque Country, Bilbao, Spain

[10] Yuki Matsumura, Atsushi Hashimoto, Shinsuke Mori, Masayuki Mukunoki, Michihiko Minoh, Clustering Scenes in Cooking Video Guided by Object Access Guidance?, 7<sup>th</sup> Workshop on Multimedia for Cooking and Eating Activities, July 3, 2015, Torino Incontra, Centro Congressi, Torino, Italy

[11] 松村優樹、橋本敦史、椋木雅之、美濃導彦、物体の把持解放を手がかりとした調理映像からの動作区間検出、2015年電子情報通信学会総大会、2015年3月11日、立命館大学びわこ・くさつキャンパス(滋賀県草津市)

[12] 山肩洋子、今堀慎治、笹田鉄郎、森信介、料理名が同じレシピ間の手順構造マッピングによる動作名オントロジーの導出、電子情報通信学会HCGシンポジウム、2014年12月19日、海峽メッセ下関(山口県下関市)

[13] Yoko Yamakata, Asuka Miyazawa, Atsushi Hashimoto, Takuya Funatomi, Michihiko Minoh, A Method for Detecting Gaze-required action while Cooking for Assisting Video Communication, Workshop on Smart Technology for Cooking and Eating Activities (in conjunction with UbiComp2014), September 14, 2014, Motif Seattle Hotel (Seattle, USA)

[14] Hiroyuki Maeta, Tetsuro Sasada, Shinsuke Mori, A Framework for Recipe Text Interpretation, Workshop on Smart Technology for Cooking and Eating Activities (in conjunction with UbiComp2014), September 14, 2014, Motif Seattle Hotel (Seattle, USA)

[15] Atsushi Hashimoto, Tetsuro Sasada, Yoko Yamakata, Shinsuke Mori, Michihiko Minoh, KUSK Dataset: Toward a Direct Understanding of Recipe Text and Human Cooking Activity, Workshop on Smart Technology for Cooking and Eating Activities (in conjunction with UbiComp2014), September 14, 2014, Motif Seattle Hotel (Seattle, USA)

[16] 山肩洋子, 今堀慎治, 前田浩邦, 森信介, 調理手順文書の自然言語解析結果からの食材・加工からなる作業ツリーの構築、データ工学研究会 (DE)、2014年9月11日、国立情報学研究所 (東京都千代田区)

[17] Atsushi Hashimoto, Jin Inoue, Takuya Funatomi, Michihiko Minoh, How does User's Access to Object Make HCI Smooth in Recipe Guidance?, HCI International 2014, June 25, 2014, Creta Maris Resort and Convention Centre (Crete, Greece)

[18] Shinsuke Mori, Graham Neubig, Language Resource Addition: Dictionary or Corpus?, The 9<sup>th</sup> International Conference on Language Resources and Evaluation (LREC), May 29, 2014, the Harpa Conference Center (Reykjavik, Iceland)

[19] Shinsuke Mori, Hirokuni Maeta, Yoko Yamakata, Tetsuro Sasada, Flow Graph Corpus from Recipe Texts, The 9<sup>th</sup> International Conference on Language Resources and Evaluation (LREC), May 28, 2014, the Harpa Conference Center (Reykjavik, Iceland)

[20] 門脇拓也, 山肩洋子, 森信介, 田中克己, 誕生・使用事由によるレシピ検索 ~ 生い立ちレシピサーチ、第6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)、2014年3月3日、淡路夢舞台&ウェスティン淡路 (兵庫県)

[21] Yoko Yamakata, Shinji Imahori, Yuichi Sugiyama, Shinsuke Mori, Katsumi Tanaka, Feature Extraction and Summarization of Recipes Using Flow Graph, The 5<sup>th</sup> International Conference on Social Informatics, November 26, 2013, 京都大学百年時計台記念館 (京都府)

[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、

情報処理学会自然言語処理研究会、2013年11月25日、屋久島環境文化村センター (鹿児島県)

[23] Ryuta Yasuoka, Atsushi Hashimoto, Takuya Funatomi, Michihiko Minoh, Detecting start and end times of object-handlings on a table by fusion of camera and load sensors, The 5<sup>th</sup> Workshop on Multimedia for Cooking and Eating Activities, October 21, 2013, Centre Convencions Internacional Barcelona, Universitat Pompeu Fabra (Barcelona, Spain)

[24] 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、電子情報通信学会データ工学研究会、2013年9月13日、国立情報学研究所 (東京都)

[25] 安岡竜太, 橋本敦史, 船富卓哉, 美濃導彦, 荷重センサを用いたインタラクション検出、電子情報通信学会食メディア研究会2013年8月研究会、2013年8月20日、公立ほこだて未来大学 (北海道)

[26] 加茂田玲奈, 上田真由美, 船富卓哉, 飯山将晃, 美濃導彦, 見え特徴と荷重バランス特徴を用いた冷蔵庫における食材管理手法、第19回画像センシングシンポジウム (SSII2013)、2013年6月13日、パシフィコ横浜 (神奈川県)

[27] 安岡竜太, 橋本敦史, 船富卓哉, 美濃導彦, カメラ荷重センサの統合による机上物体に対するハンドリング開始・終了の検出、電子情報通信学会マルチメディア・仮想環境基礎研究会、2013年3月11日、福岡工業大学 (福岡県)

[28] 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, December 13, 2012, British Computer Society, London (United Kingdom)

[29] Atsushi Hashimoto, Jin Inoue, Kazuaki Nakamura, Takuya Funatomi, Mayumi Ueda, Yoko Yamakata, Michihiko Minoh, Recognizing Ingredients at Cutting Process by Integrating Multimodal Features, The 4<sup>th</sup> Workshop on Multimedia for Cooking and Eating Activities, November 2, 2012, 奈良県新公会堂 (奈良県)

[30] 橋本敦史, 中村和晃, 船富卓哉, 美濃導彦, TexCutとパッチ型背景モデルの組み合わせによる机上物体検出システム、画像の認識・理解シンポジウム、2012年8月6日、福岡国際会議場 (福岡県)

[31] Shinsuke Mori, Tetsuro Sasada, Yoko Yamakata, Koichiro Yoshino, A Machine Learning Approach to Recipe Text Processing, Proceedings of the ECAI-Workshop on Cooking with Computers, August 3, 2012, University of Science and Letters (Montpellier, France)

[32] Rena Kamoda, Mayumi Ueda, Takuya Funatomi, Masaaki Iiyama, Michihiko Minoh, Grocery Re-identification using Load Balance Feature on the Shelf for Monitoring Grocery Inventory, Proceedings of the ECAI-Workshop on Cooking with Computers, August 3, 2012, University of Science and Letters (Montpellier, France)

[33] 宮澤飛鳥、中村和晃、橋本敦史、舩富卓哉、美濃導彦、調理者の手と容器の位置関係を利用した「かき混ぜる」行動の認識、電子情報通信学会データ工学研究会(DE)、2012年6月5日、国立情報学研究所(東京都)

〔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

## 6. 研究組織

### (1) 研究代表者

美濃導彦 (MINOH, Michihiko)  
京都大学・学術情報メディアセンター・教授  
研究者番号: 70166099

### (2) 研究分担者

棕木雅之 (MUKUNOKI, Masayuki)  
京都大学・学術情報メディアセンター・准教授  
研究者番号: 20283640  
(H24~H26)

森信介 (MORI, Shinsuke)  
京都大学・学術情報メディアセンター・准教授  
研究者番号: 90456773

山肩洋子 (YAMAKATA, Yoko)  
京都大学・情報学研究科・准教授  
研究者番号: 60423018  
(H24~H26)

舩富卓哉 (FUNATOMI, Takuya)  
京都大学・学術情報メディアセンター・助教  
(H27.3 より奈良先端科学技術大学院大学・情報科学研究科・准教授)

研究者番号: 20452310  
(H24~H26)

中村和晃 (NAKAMURA, Kazuaki)  
京都大学・法学(政治学)研究科(研究院)・助手  
研究者番号: 10584047  
(H24)

橋本 敦史 (Atsushi Hashimoto)  
京都大学・法学(政治学)研究科(研究院)・助手  
(H27 より同大・教育学研究科・助教)  
研究者番号: 80641753  
(H25~H27)

### (3) 連携研究者

中村和晃 (NAKAMURA, Kazuaki)  
大阪大学・工学(系)研究科(研究院)・助教  
研究者番号: 10584047  
(H25~H27)  
舩富卓哉 (FUNATOMI, Takuya)  
奈良先端科学技術大学院大学・情報科学研究科・准教授  
研究者番号: 20452310  
(H27)