

令和 6 年 6 月 27 日現在

機関番号：14301

研究種目：基盤研究(B)（一般）

研究期間：2020～2023

課題番号：20H04210

研究課題名（和文）手順文書からの知識獲得

研究課題名（英文）Knowledge Acquisition from Procedural Texts

研究代表者

森 信介（MORI, Shinsuke）

京都大学・学術情報メディアセンター・教授

研究者番号：90456773

交付決定額（研究期間全体）：（直接経費） 13,600,000円

研究成果の概要（和文）：自然言語理解の課題に対して知識の重要性が指摘されている。本研究では、対象を手順文書に限定し、(1) どの程度の量と質の知識が言語理解に有効か、(2) 有効であるといえる知識をどのように獲得するかという問いに取り組んだ。研究期間後半では、予定外ではあるが、知識表現として作業後の視覚的な状態を記述し、それを予測する課題を提案するとともに、解決方法の提案した。さらに、対象を Wet Lab. に拡張する取り組みを行い、知識の有用性を示すアプリケーションを構築し紹介した。

研究成果の学術的意義や社会的意義

本課題を含む一連の研究をいわゆる Wet Lab. に応用する端緒となった。本課題提案前から、レシピと調理動画はデータの取得やアノテーションが容易であることを理由に対象としていたが、研究の有用性としては Wet Lab. への応用を目指していた。本課題及びそれ以前の研究により、Wet Lab. に協力をお願いできる成果を得ていたため、大阪医科薬科大学にてその研究発表を行い、データ取得協力の合意を得たことによる。そのデータにこれまでに得た知見を用いて有用なアノテーションを行い、データセットとして公開するとともに、論文誌上で発表した。その先駆性が評価され、言語処理学会の論文賞を受賞した。

研究成果の概要（英文）：Here is the abstract of the most fruitful result of this project. In this study, we propose an egocentric biochemical video-and-language dataset called BioVL2 comprising eight videos for each of four experiments, with a total duration of 2.5 hours for all 32 samples. Each video corresponds to a protocol and two types of linguistic annotations are provided: (1) video-and-text alignment and (2) bounding boxes linked to objects in the protocol. As an application of the BioVL2 dataset, we consider the task of generating a protocol from an experimental video. Our experimental results show that the proposed system can generate better protocols than a weak baseline designed to output objects appearing in the video frames. The BioVL2 dataset will be released for research purposes only.

研究分野：自然言語処理

キーワード：言語処理 画像認識 手順書 実施動画 生化学実験 実験室DX 動作結果予測 作業支援

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

自然言語理解の研究は長期にわたって行われており、入力と出力が明確でかつ学習データが大量に与えられている課題に対しては、機械学習の進展によって高い精度が実現できている。しかしながら、学習データが大量にあるという前提は、一部の課題にのみ成り立つのみで、自然言語理解は依然として大きな課題である。この自然言語理解の課題に対して知識の重要性が指摘されている。

2. 研究の目的

本研究では、対象を手順文書に限定し、

- (1) どの程度の量と質の知識が言語理解に有効か、
- (2) 有効であるといえる知識をどのように獲得するかという問いに答えることを目的とする。

3. 研究の方法

自動獲得の正確性を人手で記述した知識との比較により評価するとともに、知識の量的十分性を以下の応用における有用性として調査する(外因評価)。

- 手順文書の意味表現の推定
- 文書検索
- 動画・写真列からの手順文書生成
- 作業支援対話システム

手順文書の例としては量・多様性ともに最大と考えられる料理レシピを最初の対象とする。しかしながら、手法としては一般性を重視し、契約書や発明開示書さらには介護記録など、客観的事実を記述した文書での有用性も評価する。さらに、獲得した知識を公開することで研究コミュニティや産業界に貢献する。

本研究提案の対象言語は、日本語と英語とする。英語に対しては、John Carroll 教授 (Sussex 大学, 英国) と Danushka Bollegala 准教授 (Liverpool 大学, 英国) から協力の承諾を得ている。

4. 研究成果

以下の研究成果を得ている。

- English Recipe Flow Graph Corpus
Yoko Yamakata, Shinsuke Mori, John Carroll
LREC, pp.5187-5194, 2020.

本研究では、本課題開始前に構築発表してあった日本語に対するレシピフローグラフの成果をもとに、英語レシピへの拡張を提案し、コーパスを公開した。

- 重要語に着目した写真列からのレシピの自動生成
西村 太一, 橋本 敦史, 森 信介
自然言語処理, Vol.27, No.2, pp.275-279, 2020.

本研究では、写真付きレシピの作成を容易にするために、写真列を入力としてレシピを生

成するという課題と、それを解決する手法を提案した。

- 手順構造を考慮した作業映像からの手順書生成

西村 太一, 橋本 敦史, 牛久 祥孝, 亀甲 博貴, 森 信介

言語処理学会第 27 回年次大会, 2021.

本研究では、フローグラフで表現される用語間の関係を写真列からの手順書生成に用いるというアイデアを動画列へ拡張し、かつ教師なしで手順構造をモデル化することで、一貫性のある手順書を生成する手法を提案し、有効性を実験的にしめした。

- レシピ分野における動作対象の状態変化を考慮したデータセットの構築と検索モデルの提案

白井 圭佑, 橋本 敦史, 牛久 祥孝, 栗田 修平, 亀甲 博貴, 森 信介

言語処理学会第 28 回年次大会講演論文集, G1-1, 2022.

この研究では、本課題の当初の計画を発展させ、言語理解を画像検索で評価するために、新たなタスクを提案した。調理動作を対象に動作後の状態として適切な画像を検索することを言語理解とした。結果的には、この後の各所の研究が画像言語の融合に向かったので、本研究はすこしではあるが先駆的であったといえる。

- BioVL2 データセット: 生化学分野における一人称視点の実験映像への言語アノテーション

西村 太一, 迫田 航次郎, 牛久 敦, 橋本 敦史, 奥田 奈津子, 小野 富三人, 亀甲 博貴, 森 信介

自然言語処理, 2022, No. 29, Vol. 4. (2022 年言語処理学会論文賞)

この論文は、本課題を含む一連の研究をいわゆる Wet Lab. に応用する端緒となっている。本課題提案前から、レシピと調理動画はデータの取得やアノテーションが容易であることを理由に対象としていたが、研究の有用性としては Wet Lab. への応用を目指していた。本課題及びそれ以前の研究により、Wet Lab. に協力をお願いできる成果を得ていたため、大阪医科薬科大学にてその研究発表を行い、データ取得協力の合意を得たことによる。そのデータにこれまでに得た知見を用いて有用なアノテーションを行い、データセットとして公開するとともに、論文誌上で発表した。その先駆性が評価され、言語処理学会の論文賞を受賞した。

- 調理動作後の物体の視覚的状态予測を目指した Visual Recipe Flow データセットの構築と評価

白井 圭佑, 橋本 敦史, 西村 太一, 亀甲 博貴, 栗田 修平, 森 信介

自然言語処理, No. 30, Vol. 3., 2023

本研究では、知識の表現としてフローグラフのノードに中間状態の写真を付与したデータを提案し、それを構築するとともに、それを用いた手順理解の評価を行った。本課題の当初の計画から発展し、言語内での記号的な知識のみならず、動作後の視覚的な状態を記述し、それを予測する課題を提案するとともに、解決方法の提案に発展している。

- 一人称視点映像を用いたマルチモーダル作業支援システム

梶村 恵矢, 西村 太一, 羽路 悠斗, 山本 航輝, 崔 泰毓, 亀甲 博貴, 森 信介

言語処理学会第 30 回年次大会, 2024.

本発表では、本課題の応用先として Wet Lab. などでの知識の有用性を示すアプリケーションを構築し紹介した。本課題の終了の後に、Wet Lab. などへの社会実装を進めており、このアプリケーションはそのプロトタイプである。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件 / うち国際共著 0件 / うちオープンアクセス 2件）

1. 著者名 Taichi Nishimura, Atsushi Hashimoto, Yoshitaka Ushiku, Hiroataka Kameko, Shinsuke Mori	4. 巻 82
2. 論文標題 State-aware Video Procedural Captioning	5. 発行年 2023年
3. 雑誌名 Multimedia Tools and Applications	6. 最初と最後の頁 37273-37301
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Yixin Zhang, Yoko Yamakata, and Keishi Tajima	4. 巻 16
2. 論文標題 MIRecipe: A Recipe Dataset for Stage-Aware Recognition of Changes in Appearance of Ingredients	5. 発行年 2021年
3. 雑誌名 ACM Multimedia Asia 2021	6. 最初と最後の頁 1-7
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Akihisa Ishino, Yoko Yamakata, Hiroaki Karasawa and Kiyoharu Aizawa	4. 巻 Technical Demo
2. 論文標題 RecipeLog: Recipe Authoring App for Accurate Food Recording	5. 発行年 2021年
3. 雑誌名 ACM Multimedia 2021	6. 最初と最後の頁 2798-2800
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 西村 太一, 橋本 敦史, 森 信介	4. 巻 27
2. 論文標題 重要語に着目した写真列からのレシピの自動生成	5. 発行年 2020年
3. 雑誌名 自然言語処理	6. 最初と最後の頁 275-279
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計23件（うち招待講演 2件 / うち国際学会 10件）

1. 発表者名 Taichi Nishimura, Atsushi Hashimoto, Yoshitaka Ushiku, Hirotaka Kameko, Shinsuke Mori
2. 発表標題 State-aware Video Procedural Captioning
3. 学会等名 ACM Multimedia (国際学会)
4. 発表年 2023年

1. 発表者名 Taichi Nishimura, Kojiro Sakoda, Atsushi Hashimoto, Yoshitaka Ushiku, Natsuko Tanaka, Fumihito Ono, Hirotaka Kameko, Shinsuke Mori
2. 発表標題 Egocentric Biochemical Video-and-Language Dataset
3. 学会等名 CLVL (ICCV workshop) (国際学会)
4. 発表年 2021年

1. 発表者名 Taichi Nishimura, Suzushi Tomori, Hayato Hashimoto, Atsushi Hashimoto, Yoko Yamakata, Jun Harashima, Yoshitaka Ushiku, Shinsuke Mori
2. 発表標題 Visual Grounding Annotation of Recipe Flow Graph
3. 学会等名 LREC2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Yoko Yamakata, Shinsuke Mori, John Carroll
2. 発表標題 English Recipe Flow Graph Corpus
3. 学会等名 LREC2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Keisuke Shirai, Hirotaka Kameko, Shinsuke Mori
2. 発表標題 Towards Flow Graph Prediction of Open-Domain Procedural Texts
3. 学会等名 CL, The 8th Workshop on Representation Learning for NLP (国際学会)
4. 発表年 2023年

1. 発表者名 Satayu Parinayok, Yoko Yamakata, and Kiyoharu Aizawa
2. 発表標題 Open-Vocabulary Segmentation Approach for Transformer-Based Food Nutrient Estimation
3. 学会等名 ACM International Conference on Multimedia in Asia (国際学会)
4. 発表年 2023年

1. 発表者名 Liangyu Wang, Yoko Yamakata, and Kiyoharu Aizawa
2. 発表標題 Automatic Dataset Creation from User-generated Recipes for Ingredient-centric Food Image Analysis
3. 学会等名 ACM International Conference on Multimedia in Asia (国際学会)
4. 発表年 2023年

1. 発表者名 Nanba, H., Kubo, S., and Fukuda, S
2. 発表標題 Automatic Generation of Explanatory Text from Flowchart Images in Patents
3. 学会等名 Workshop on Patent Text Mining and Semantic Technologies
4. 発表年 2023年

1. 発表者名 白井 圭佑, 亀甲 博貴, 森 信介
2. 発表標題 オープンメインの手順書のフローグラフ予測とデータセットの構築
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2023年

1. 発表者名 Taichi Nishimura, Atsushi Hashimoto, Yoshitaka Ushiku, Hirotaka Kameko, Shinsuke Mori
2. 発表標題 Recipe Generation from Unsegmented Cooking Videos
3. 学会等名 arXiv (国際学会)
4. 発表年 2022年

1. 発表者名 Yoko Yamakata, Akihisa Ishino, Akiko Sunto, Sosuke Amano, Kiyoharu Aizawa
2. 発表標題 Recipe-oriented Food Logging for Nutritional Management
3. 学会等名 ACM Multimedia (国際学会)
4. 発表年 2022年

1. 発表者名 Yoko Yamakata
2. 発表標題 Food Computing for Human and Planet Health
3. 学会等名 INTERNATIONAL WORKSHOP ON ADVANCED IMAGE TECHNOLOGY (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 山肩洋子, 王良羽, 山末英嗣, 駿藤晶子, 相澤清晴
2. 発表標題 ユーザ生成型レシピからの栄養価および環境影響の自動計算のための食材名分類モデルの構築
3. 学会等名 電子情報通信学会 技術研究報告
4. 発表年 2022年

1. 発表者名 佐藤賢志郎, 山肩洋子, 相澤清晴
2. 発表標題 一人称視点の食事映像からの食行動レポートの自動生成
3. 学会等名 電子情報通信学会 技術研究報告
4. 発表年 2022年

1. 発表者名 難波英嗣, 久保翔平, 福田悟志
2. 発表標題 特許中のフローチャート画像からの説明文の自動生成
3. 学会等名 情報処理学会 第150回 情報基礎とアクセス技術研究発表会
4. 発表年 2023年

1. 発表者名 難波英嗣
2. 発表標題 自然言語処理技術を利用した特許文献の分析
3. 学会等名 技術情報協会 (招待講演)
4. 発表年 2022年

1. 発表者名 石野耀久、山肩洋子、唐澤弘明、天野宗佑、相澤清晴
2. 発表標題 RecipeLog: 手順構造を持ったマルチメディアレシピのための入力インタフェース
3. 学会等名 映像情報メディア学会 メディア工学研究会 映情学技報
4. 発表年 2022年

1. 発表者名 佐藤賢志郎、天野宗佑、山肩洋子、相澤清晴
2. 発表標題 食行動の自動認識に向けた食事の一人称視点映像データセットの構築
3. 学会等名 映像情報メディア学会 メディア工学研究会 映情学技報
4. 発表年 2022年

1. 発表者名 樊 エイブン, 福田 悟志, 難波 英嗣,
2. 発表標題 特許中の画像とテキストを用いた手順オントロジーの構築
3. 学会等名 情報処理学会第145回情報基礎とアクセス技術研究発表会
4. 発表年 2022年

1. 発表者名 西村 太一, 橋本 敦史, 牛久 祥孝, 森 信介
2. 発表標題 映像からのストーリー生成: イベント選択器と文生成器の同時学習
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2022年

1. 発表者名 白井圭佑, 橋本敦史, 牛久祥孝, 栗田修平, 亀甲博貴, 森信介
2. 発表標題 レシビ分野における動作対象の状態変化を考慮したデータセットの構築と検索モデルの提案
3. 学会等名 言語処理学会第28回年次大会
4. 発表年 2022年

1. 発表者名 西村 太一, 友利 涼, 橋本 隼人, 橋本 敦史, 山肩 洋子, 原島 純, 牛久 祥孝, 森 信介
2. 発表標題 レシビフローグラフへのVisual Groundingアノテーション
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 西村 太一, 橋本 敦史, 牛久 祥孝, 森 信介
2. 発表標題 写真列と構造要素からの手順構造と手順書の同時学習
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

大規模テキストアーカイブ研究分野 http://www.lsta.media.kyoto-u.ac.jp/ フローグラフコーパス (FG corpus) http://www.lsta.media.kyoto-u.ac.jp/resource/data/recipe/ John A. Carroll Publications http://users.sussex.ac.uk/~johnca/publications.html

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	難波 英嗣 (NAMBA Hidetsugu) (50345378)	中央大学・理工学部・教授 (32641)	
研究分担者	亀甲 博貴 (KAMEKO Hirotaka) (50827524)	京都大学・学術情報メディアセンター・助教 (14301)	
研究分担者	山肩 洋子 (YAMAKATA Yoko) (60423018)	東京大学・大学院情報理工学系研究科・准教授 (12601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
英国	University of Sussex			