

令和 4 年 6 月 9 日現在

機関番号：32689

研究種目：基盤研究(B) (一般)

研究期間：2017～2020

課題番号：17H01831

研究課題名(和文) 動画に対する深い意味的注釈の付与

研究課題名(英文) Deep semantic annotation of video contents

研究代表者

林 良彦 (Hayashi, Yoshihiko)

早稲田大学・理工学術院・教授(任期付)

研究者番号：80379156

交付決定額(研究期間全体)：(直接経費) 12,500,000円

研究成果の概要(和文)：動画コンテンツにおいて描写されている人間の動作・行為に対して、その意味を考慮した適切な注釈付けを行うことができれば、高度な動画検索システムを実現したり、動画から有用な知識情報を計算機が利用可能な形で抽出することが可能となる。本研究では、そのための基盤技術として、フレーム画像の内容を離散的構造として抽出するためのシーングラフ生成法、および、自然言語表現として得るための動作キャプション生成法の研究を進め、比較手法を上回る精度を達成した。また、関連する画像・言語解析技術の基礎研究を推進し、その成果を論文として発表した。

研究成果の学術的意義や社会的意義

動画コンテンツは継続的に増加しており、これらのコンテンツに対する高度なアクセス手段の提供や、コンテンツから有用な情報や知識を抽出機能が求められている。本研究の成果は、このような機能・システムを実現するための基盤技術として活用することが可能である。開発した技術は、映像・画像情報と言語情報を統合的に利用するものであり、マルチメディア・マルチモーダル情報処理技術の高度化に寄与する。

研究成果の概要(英文)：To enable an advanced retrieval system or an intelligent knowledge extraction system that deals with a large set of video contents, it is essential to semantically annotate them adequately. Towards this ultimate goal, this study researched fundamental technologies that combine vision and language technologies. More specifically, we have developed an effective yet efficient scene graph generation systems and an action captioning system. Empirical results show that the resulting systems generally performed better than the comparative systems. These systems respectively achieve information structure adequate for computer processing and for human consumption.

研究分野：自然言語処理

キーワード：情報資源の構築・管理 動画 意味的注釈 シーングラフ生成 キャプション生成 動詞意味論 語彙  
意味関係 オントロジー

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。

## 1. 研究開始当初の背景

本研究課題「動画に対する深い意味的注釈の付与」の開始当初(2017年度)においては、すでに多くの動画コンテンツが利用可能となっていたが、動画の内容に基づく効率的なアクセス手段、特に、動画に描写されている動作や行為などに対する意味的な検索・アクセスを行う手段の実現が望まれていた。一方で、これに関連する画像・動画解析や自然言語処理技術は、深層学習の発展に伴い、実用的な精度を達成しうる要素技術が提案されつつあった。例えば、画像中から物体を検出する際の代表的な手法である Faster R-CNN や、自然言語文を扱う際の基本技術となっている Transformer などが登場してきており、その有効性も明らかになってきていた。このため、これらの技術を利用することにより、動画に対して、動画に描写されている動作や行為などに関する意味的注釈を与えることにより、意味的な検索・アクセスを行うシステムの実現へ向けた研究が実施可能であると考えた。

## 2. 研究の目的

本研究課題「動画に対する深い意味的注釈の付与」は、主に動画において描写されているエージェント(人や動物など)の動作や行為に対して、意味的な注釈を付与することを目的としている。上述の技術的背景を考慮し、この目的へ向け、本研究では以下のような技術目標を達成することを目的とした。

- (1) フレーム画像からのシーングラフ生成: 動画におけるフレーム画像に描写されている物体、および、それらの間の関係をシーングラフと呼ぶグラフ構造として抽出し、これを動作や行為の時間的な変化に沿って展開するために必要な技術要素を明らかにする。
- (2) 動作・行為に対するキャプション生成: シーングラフは計算機処理に適した離散的情報構造であるが、より人間に理解しやすい注釈として、動画において描写されている動作・行為を端的に表す自然言語文=キャプションを生成するための方式を明らかにする。
- (3) 動画検索システムの実現と評価: 大規模な動画コンテンツ集合から、検索クエリに意味的に適合する動画区間を検出するための動画検索システムの構成方式を明らかにするとともに、上述の意味的注釈を検索において利用する方法を調査する。
- (4) 関連技術に関する基礎的・応用的研究: 動画の利用、基本的な自然言語解析に関する基礎的・応用的研究を実施し、本研究の目的をより効率的・効果的に達成するためのフィードバックを得る。

## 3. 研究の方法

- (1) フレーム画像からのシーングラフ生成: 一般に画像中には多くの物体が描写されているため、物体およびそれらの間の関係を構造化するシーングラフ生成においては、計算量が課題となる。そこで、言語情報や世界知識に関する情報を利用することにより、効率的な生成手法を検討することとした。また、シーングラフの時間展開を行うために必要と考えられる動詞の意味論に関して基礎的な研究を進めることとした。
- (2) 動作・行為に対するキャプション生成: すでに動画からキャプションを生成する技術の研究が進んでいたことから、この技術を利用して一般的なキャプション文をまず生成し、それを意味注釈に適した形の動作記述の系列へと変換する方法を探究することとした。また、エージェントがロボットであるという将来的な状況も考慮し、ロボットの物理動作と動作の言語記述の間の相互変換に関する研究も行うこととした。
- (3) 動画検索システムの実現と評価: 動画検索に関する shared task (具体的には TRECVID AVS: Ad-hoc Video Search) に参加することにより、広い見地からの技術評価を行うこととした。
- (4) 関連技術に関する基礎的・応用的研究: 動画コンテンツの利用に関しては、動画を用いた映像監視システムの研究に参加し、本研究に関連する技術をフィードバックすることとした。また、基本的な自然言語解析に関しては、シーングラフ生成などにおいて制約として利用できる可能性のある意味属性・意味関係知識やコモンセンス知識の利用に関する基礎的な研究を実施することとした。

## 4. 研究成果

### (1) フレーム画像からのシーングラフ生成

シーングラフ(図1)とは、画像に含まれる物体とそれらの間の関係を記述するグラフである。動画から抽出したフレーム画像に対して適切なシーングラフを生成できれば、動画に対する意味的注釈の要素として利用することが可能となる。シーングラフの生成においては、描写されている物体とそれらの間の関係を三つ組の集合として表すことが必要とな



図 1. シーングラフの例 [1]

る。意味的かつ世界的に適合したシーングラフを効率的に実現することが必要となる。この問題に対して、主に以下の2つの方法によりアプローチした。

**言語の意味的特徴量の利用 [1]:** 画像中の物体のクラスを離散的な記号ではなく、類似度が定義可能な連続的な意味的特徴量により表現することにより、学習データに存在しない未知の物体が認識時に現れた場合にも、類似する既知の物体と同様に扱うことが可能とした。また、連続的な言語特徴量の活用を可能とするため、CRF とニューラルネットワークを組み合わせたアーキテクチャを提案した。多重線型な層を組み込むことにより、計算量を抑えつつ、高精度なシーングラフ生成を実現した。物体を意味的特徴量により表現することにより、意味注釈の計算機による利用範囲が広がることも期待できる。

**シーンにおける階層的な構造の探索と利用 [2]:** シーングラフは局所的な物体間の関係を表す部分グラフが結合した情報構造である。ここで一つの課題として、シーングラフ全体にわたる大局的な意味的整合性を保つことがあげられる。このために、シーンにおける物体の階層関係 (図2) を探索し、さらにそれを制約として利用することを提案した。局所的・大局的な情報を相互に干渉させることにより全

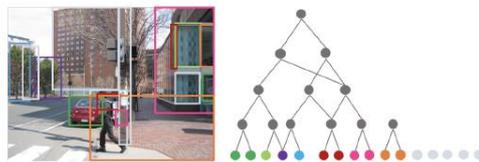


図 2. 物体間の階層関係 [1]

体最適なグラフ構成を生成可能とするアーキテクチャ (図3) により、さらに高精度なシーングラフを生成することが可能となった。なお、本方式の検討を通して、実世界に関する常識的な知識 (コモンセンス知識) を利用することの有用性についての感触を得た。

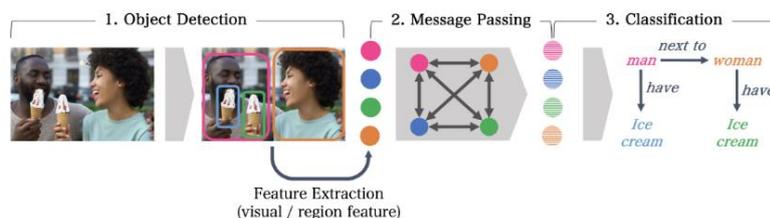


図 3. シーングラフ生成システムの構成 [1]

## (2) 動作・行為に関するキャプション生成

シーングラフは計算機処理に適した離散的構造であるが、自然言語による記述 (キャプション) は、より人間にもわかりやすい注釈である。また、自然言語解析を行って適切な構造化を施すことにより、検索エンジンにおけるインデックス情報としても利用可能である。そこで、動作・行為に関するキャプション (以下、動作記述と呼ぶ) 生成の課題に取り組んだ [3, 4]。

動作記述は、人間が動作を行っている動画中の区間に対して、その動作を簡潔に記述するものであり、(a) 独立した単文の系列、(b) 各単文は人間が主語で、動詞は現在形もしくは現在進行形、という言語特徴を持つものとした。また、近年の動画キャプション生成技術の進展を考慮し、一気に動作記述の生成を目指すのではなく、既存技術により生成されたキャプションを上記の要件を満たす文系列へと変換する方式を提案した (図4)。

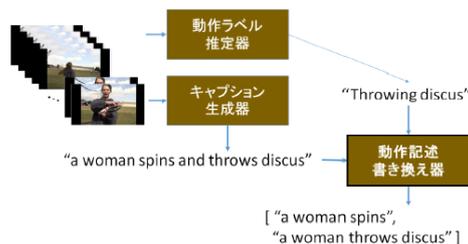


図 4. 書き換え処理による動作記述生成 [4]

書き換え処理は、翻訳処理や要約処理においても利用されている sequence-to-sequence により行うが、キャプションから動作記述への変換という当該タスクの特性に基づき、教師データを作成し学習を行った。また、対象の動作の大きな分類を表す動作ラベルを用いることにより、書き換えの精度向上を狙った。評価実験の結果からは、動作ラベルを用いることにより、漏れや重複の少ない書換を効率よく得ることができることを確認した。

さらに、得られたキャプションが動画検索におけるインデックス情報として適切に利用可能であるかを数種類の検索アルゴリズムを用いて評価した。その結果、最新の文のエンコーディング手法である BERT を用いると、動作ラベルを利用した動作記述によって、高い検索精度が得られることが確認できた。この結果は、提案手法により、簡潔で端的な動作記述を得ることができること、最新の言語処理技術と組み合わせることにより、高度な検索利用への道が開かれることを示しており、今後のマルチメディア検索技術へ貢献しうる成果であると考えられる。

ここまで述べた研究は、すでに作成された動画コンテンツを対象としたものである。動作を対象としたキャプション生成は、ロボットの振る舞いを描写した動画においても重要であると考えられる。また、ロボットの振る舞いを直接的に表す時系列データ (典型的には数か所のロボットの関節角度の情報) との双方向変換が可能となれば、ロボットの動作状況の説明

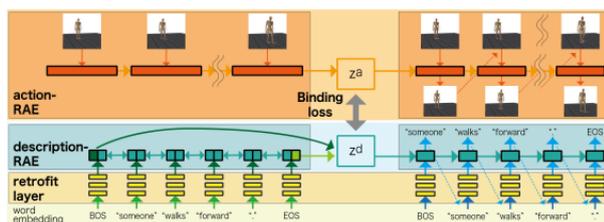


図 5. 動作と言語記述の相互変換機構 [5]

や、自然言語記述によるロボットへの動作指示が可能となる。以上のような考えのもと、ロボットのシミュレーションデータを対象に、動作と言語の双方向変換の研究も行った [5, 6]。図 5 にシステム構成を示す。シーングラフ生成においても用いた意味的特徴量を用いることにより、動作との対応付けが学習されていない単語も利用可能であり、さらに、実世界の動作情報を用いた学習を行うことにより、テキストだけからでは得られない特徴表現が得られることなどが明らかとなった。

### (3) 動画検索システムの実現と評価

上述の動画検索実験は、生成された動作記述の基本的な有効性を確認するものであった。より実用に近いセッティングにおける関連技術の有効性を評価し、技術的な課題を抽出するため、本研究チームは、大規模な動画検索のコンペティションである TRECVID AVS (Ad-hoc Video Search) に継続的に参加し、良好な成績を収めてきている [7]。以下では、語義曖昧性解消、あるいは、フィルタリグという自然言語処理における基本技術が TRECVID AVS のようなタスクにおいても有効であることを示した結果 [8] について述べる。まず、図 6 に検索システムの構成を示す。本システムの基本的な考え方は、動画検索において有用と考えられる多数の物体をそれぞれ識別しうる識別器を事前に多数用意しておくということである。このため、検索精度の向上のためには、適切な識別を起動することが重要である。

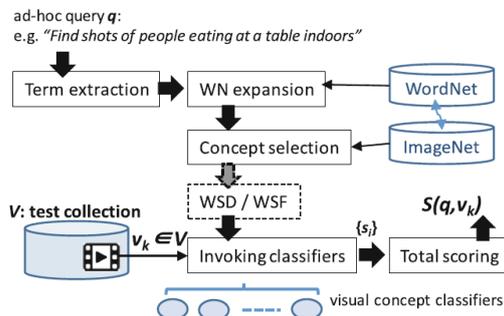


図 6. AVS システムの構成 [8]

システムはまず、自然言語文で記述されたクエリから検索に有用な語を抽出 (term extraction) し、それを語彙資源 WordNet を用いて語義展開し (WN expansion)、画像可能かにより選別する (concept selection)。一般に多くの語は多義であるため、適切な語義曖昧性解消 (word sense disambiguation: WSD)、もしくは、無関係な語義の絞り込み (word sense filtering: WSF) を行うことにより、適切な画像識別器を選択することが可能となり、これは検索精度の向上に貢献することが期待できる。

そこで本研究では、クエリ中のキーワードの語義曖昧性の解消・絞り込みを語義・概念の分散表現を用いることにより行う手法を数種類実装し、その精度が最終的な動画検索の精度にどのように影響するかを調べた。その結果、平均的な語義解消精度と精度は中程度の相関であることが分かり、語義解消精度のさらなる向上の必要性が明らかとなった。一方で、複数の語義を許容する語義絞り込みにより、動画検索の精度向上が見られた。以上から、動画検索の精度向上のためにはクエリ中のキーワードの適度な語義絞り込みが有効であることが確認できた。

### (4) 関連技術に関する基礎的・応用的研究

本研究チームでは、動画の利用、基本的な自然言語解析に関するさまざまな基礎的・応用的研究を実施してきている。以下では、代表的な研究例について述べる。

#### 自然言語解析に関する基礎的研究:

- **動詞語義の分類 [9]:** 動作を言語で記述する上で最も基本となるのは動詞である。名詞概念と同様に動詞で表される概念についても上下関係 (包摂関係) が考えられるが、この扱いは語彙資源によって異なる。本研究では日本語の動詞を対象とし、EDR 電子化辞書、述語項構造ソーラスを対象とし、特に自動詞・他動詞の対立という意味関係の分析を進めた。行為に関する動詞に関する分析結果の一部を表 1 に示す。このような言語学的に深い分析は、より意味的に適正な深い意味注釈の付与につながる重要な研究であると考えられる。
- **意味の定義の利用と生成 [10]:** 分布意味論によれば、言語の構成要素 (単語など) の意味は、その使用の状況から機能的に獲得すべきであり、これが近年の自然言語処理の本流を推進する動向である。しかしながら、辞書における語義定義文を適切に利用できれば、より選択的な表現獲得が可能となると考えられる。本研究では、辞書定義文において暗黙的に想定される意味関係をパターンに基づいて探索し、単語のペアに対して適切な表現を獲得することに成功した。さらに、獲得された表現をもとに語義定義文を生成できることを確認した。この成果は、動画などの対象コンテンツから直接抽出できる意味内容だけではなく、辞書などの外部資源に存在する意味・知識情報を利用する可能性を開く有意義な研究であると考えられる。
- **単語の属性情報の推定 [11]:** 単語概念が持つ情報は、概念的な上位下位といった階層的なものだけでなく、単語が指す事物が有しうる属性的な情報も考えられる。本研究では、既知の単語概念が想起しうる属性情報を学習しておくことにより、未知の単語に対しても一定の精度で属性情報を推定できることを示した。また、得られた属性情報を利用することによって、

表 1. 行為の動詞における語義対応 [9]

自動詞語義分類	他動詞語義分類 (〈行為〉の下位概念)
現象 (34)	〈変える〉 → 〈ものを対象とする行為〉 → 〈対象行為〉 〈物の性質を変化させる行為〉 → 〈対象活動〉 → 〈対象行為〉 〈物に力を加える行為〉 → 〈対象活動〉 → 〈対象行為〉 〈情報の移動〉 → 〈ものを対象とする行為〉 → 〈対象行為〉 〈変える〉 → 〈ものを対象とする行為〉 → 〈対象行為〉
変化 (33)	〈行為や物事の進行にかかわる活動〉 〈覆う〉 → 〈対象活動〉 → 〈対象行為〉 〈ものとの関係にかかわる行為〉 → 〈対象行為〉 〈具体物を移動する〉 → 〈対象活動〉 → 〈対象行為〉
移動 (28)	〈物に力を加える行為〉 → 〈対象活動〉 → 〈対象行為〉 〈情報の移動〉 → 〈ものを対象とする行為〉 → 〈対象行為〉 〈具体物を移動する〉 → 〈対象活動〉 → 〈対象行為〉
状態 (21)	〈行為や物事の進行にかかわる活動〉 〈変える〉 → 〈ものを対象とする行為〉 → 〈対象行為〉 〈情報を移動する〉 → 〈対象活動〉 → 〈対象行為〉
行為 (17)	〈声を与えること〉 → 〈対人行為〉 → 〈対象行為〉 〈させる〉 → 〈対人行為〉 → 〈対象行為〉

単語間の類似度・関連度の精度を向上できることを示した。単語が指示する事物が有する属性情報は、コンテンツからの豊かな意味注釈の獲得において重要性・有用性が増すものと考えられ、本研究はこれに寄与するものと考えられる。

- **コモンセンス知識の利用:** シーングラフの生成において端的であるが、多くの可能性の中から意味的に整合した表現を得るためには、人間がもつ常識的知識（コモンセンス知識）を適切に利用することが有用であると考えられる。このようなコモンセンス知識の有用性は物語理解・生成や質問応答などの自然言語処理タスクにおいても再認識されており、近年、研究が盛んに行われている。これらの研究においては、コモンセンス知識を言語情報から生成することも行われているが、人手により構築された知識リソースを利用しようとする研究が主流である。リソースとしては、人手により構築された **ConceptNet** が代表的である。**ConceptNet** は、概念をノードとし概念間の関係をエッジとするグラフ構造であり、シーングラフと親和性の高い構造となっている。ここで、概念ノードは単語のラベルを持つが、その単語の多義性が区分されていないため、例えば、グラフ上の経路を利用するにはノードにおける「意味的ギャップ」を考慮する必要がある。本研究においては、この問題を取り上げ、教師あり学習により意味的ギャップを検出する研究の取り組みを開始した [12]。小規模ではあるが、人手によるギャップに関する評定を付与するデータセットを構築し、数種のベースライン手法を比較した。この課題は、コンテンツへの意味注釈付与において、既存のコモンセンス知識リソースを適正に利用するために不可欠と考えられる重要である。
- **動画の利用に関する応用的研究:** 本チームでは本研究課題に関するフィードバックやヒントを得ることを目的に、動画を用いた映像監視システムの研究に参加してきた。具体的には、牛の分娩予兆を検出し、ユーザ（畜産農家）にアラートを行うシステム（図7）を研究開発している [13]。主な結果として、完全に **end-to-end** でシステムを構成するのではなく、一度、特徴抽出を起こった後に、分娩予兆の検出を行ったほうが良いこと、また、特徴抽出においては、関連する特徴量をマルチタスク学習（首と尾の位置の推定、姿勢の推定）することが有効であることが確認できた。現在、深層学習に基づき、**end-to-end** で目的システムを構築することが主流となっているが、目的タスクの状況において適切な役割分担を踏まえたシステム構成とすべきことが示唆される。これは、動画に対する意味注釈付与・検索システムの構築においても有用な知見であると考えられる。

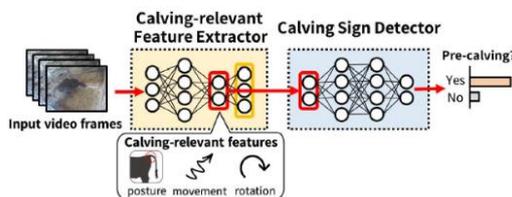


図 7. 牛の分娩予兆検出システムの構成 [13]

#### [参考文献]

1. 黒澤, 菊池, 小林, 林: 言語特徴量を利用したシーングラフ生成の効率的な計算機構. NLP2018.
2. I. Kurosawa, T. Kobayashi, Y. Hayashi: Exploring and Exploiting the Hierarchical Structure of a Scene for Scene Graph Generation. ICPR2020 (査読有).
3. 平川, 小林, 林: 動画からの動作キャプション生成における書き換え技術の適用. NLP2019.
4. 平川, 小林, 林: 動画からの動作記述の生成とその評価. NLP2020.
5. M. Toyoda, K. Suzuki, H. Mori, Y. Hayashi, T. Ogata: Embodying Pre-trained Word Embeddings through Robot Actions. ICRA2021 (査読有).
6. 豊田, 林, 鈴木, 尾形: 言語と動作の統合表現獲得による双方向変換. NLP2022.
7. K. Ueki, et al.: Waseda\_Meisei at TRECVID 2018: Ad-hoc Video Search. TRECVID2018.
8. K. Hirakawa, K. Kikuchi, K. Ueki, T. Kobayashi, Y. Hayashi: Ad-hoc Video Search Improved by the Word Sense Filtering of Query Terms. AIRS2018 (査読有).
9. 加藤: 動词语義の階層分類に関する一考察. NLP2018.
10. K. Washio, S. Sekine, T. Kato: Bridging the Defined and Defining: Exploiting Implicit Lexical Semantic Relations in Definition Modeling. EMNLP2019 (査読有).
11. M. Hasegawa, T. Kobayashi, Y. Hayashi: Word Attribute Prediction Enhanced by Lexical Entailment Tasks, LREC2020 (査読有).
12. Y. Hayashi: Towards the Detection of a Semantic Gap in the Chain of Commonsense Knowledge Triples. LREC2022 (査読有).
13. R. Hyodo, T. Nakano, T. Ogawa: Feature Representation Learning for Calving Detection of Cows using Video Frames. ICPR2020 (査読有).

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 3件）

1. 著者名 UEKI Kazuya, HIRAKAWA Koji, KIKUCHI Kotaro, KOBAYASHI Tetsunori	4. 巻 84
2. 論文標題 Zero-Shot Video Retrieval from a Query Phrase Including Multiple Concepts ?Efforts and Challenges in TRECVID AVS Task?	5. 発行年 2018年
3. 雑誌名 Journal of the Japan Society for Precision Engineering	6. 最初と最後の頁 983 ~ 990
掲載論文のDOI（デジタルオブジェクト識別子） 10.2493/jjspe.84.983	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 林 良彦	4. 巻 32
2. 論文標題 言語学とAI	5. 発行年 2017年
3. 雑誌名 人工知能学会誌	6. 最初と最後の頁 384--393
掲載論文のDOI（デジタルオブジェクト識別子） 10.11517/jjsai.32.3_384	査読の有無 無
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 Toyoda Minori, Suzuki Kanata, Mori Hiroki, Hayashi Yoshihiko, Ogata Tetsuya	4. 巻 6
2. 論文標題 Embodying Pre-Trained Word Embeddings Through Robot Actions	5. 発行年 2021年
3. 雑誌名 IEEE Robotics and Automation Letters	6. 最初と最後の頁 4225 ~ 4232
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/LRA.2021.3067862	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64件（うち招待講演 1件／うち国際学会 26件）

1. 発表者名 Koki Washio, Satoshi Sekine, and Tsuneaki Kato
2. 発表標題 Bridging the Defined and the Defining: Exploiting Implicit Lexical Semantic Relations in Definition Modeling
3. 学会等名 EMNLP2019（査読有，会議録有）（国際学会）
4. 発表年 2019年

1. 発表者名 Kazuya Ueki, Takayuki Hori, and Tetsunori Kobayashi
2. 発表標題 Waseda_Meisei_SoftBank at TRECVID 2019: Ad-hoc Video Search
3. 学会等名 TRECVID2019 (会議録有) (国際学会)
4. 発表年 2019年

1. 発表者名 Kazuya Ueki and Takayuki Hori
2. 発表標題 Zero-shot Video Retrieval using a Large-scale Video Database
3. 学会等名 MDACL@ICCV2019 (査読有, 会議録有) (国際学会)
4. 発表年 2019年

1. 発表者名 田辺ひかり, 小川哲司, 小林哲則, 林良彦
2. 発表標題 感情推定における感情カテゴリに関する先験的知識の利用
3. 学会等名 言語処理学会第26回年次大会 (NLP2020) (会議録有)
4. 発表年 2020年

1. 発表者名 平川幸司, 小林哲則, 林良彦
2. 発表標題 動画からの動作記述の生成とその評価
3. 学会等名 言語処理学会第26回年次大会 (NLP2020) (会議録有)
4. 発表年 2020年

1. 発表者名 植木 一也, 堀 隆之, 小林 哲則
2. 発表標題 クエリ文を用いた映像検索手法の比較検証
3. 学会等名 第22回 画像の認識・理解シンポジウム (MIRU2019) (会議録有)
4. 発表年 2019年

1. 発表者名 武藤 良, 桑原 明栄子, 植木 一也
2. 発表標題 マルチタスク学習を用いた顔の複数属性推定
3. 学会等名 第22回 画像の認識・理解シンポジウム (MIRU2019) (会議録有)
4. 発表年 2019年

1. 発表者名 セイエドネシャド ロスタム, 武藤 良, 堀 隆之, 植木 一也
2. 発表標題 複数の人物属性を利用した大規模映像からの人物検索
3. 学会等名 第22回 画像の認識・理解シンポジウム (MIRU2019) (会議録有)
4. 発表年 2019年

1. 発表者名 兵頭亮介, 菅原一真, 中野鐵兵, 赤羽誠, 小林哲則, 小川哲司
2. 発表標題 動画像から得られる牛の身体情報に基づく分娩予兆検知システム
3. 学会等名 電子情報通信学会 PMRU2019-1 (会議録有)
4. 発表年 2019年

1. 発表者名 兵頭亮介, 斎藤奨, 中野鐵兵, 赤羽誠, 小林哲則, 小川哲司
2. 発表標題 画像から得られる牛の身体情報に基づく分娩予兆検知
3. 学会等名 2019年度人工知能学会全国大会 (会議録有)
4. 発表年 2019年

1. 発表者名 Mika Hasegawa, Testunori Kobayashi, Yoshihiko Hayashi
2. 発表標題 Social Image Tags as a Source of Word Embeddings: A Task-oriented Evaluation
3. 学会等名 Proc. of LREC2018, pp.969-973 (国際学会)
4. 発表年 2018年

1. 発表者名 Kouji Hirakawa, Kotaro Kikuchi, Kazuya Ueki, Tetsunori Kobayashi, Yoshihiko Hayashi
2. 発表標題 Ad-hoc Video Search Improved by the Word Sense Filtering of Query Terms
3. 学会等名 Proc. of AIRS2018, pp.157-163 (国際学会)
4. 発表年 2018年

1. 発表者名 Koki Washio and Tsuneaki Kato
2. 発表標題 Undersampling Improves Hypernymy Prototypicality Learning
3. 学会等名 Proc. of LREC2018, pp.4550-4554 (国際学会)
4. 発表年 2018年

1 . 発表者名 Koki Washio and Tsuneaki Kato
2 . 発表標題 Filling Missing Paths: Modeling Co-occurrences of Word Pairs and Dependency Paths for Recognizing Lexical Semantic Relations
3 . 学会等名 Proc. of NAACL2018, pp.1123 -1133 ( 国際学会 )
4 . 発表年 2018年

1 . 発表者名 Koki Washio and Tsuneaki Kato
2 . 発表標題 Neural Latent Relational Analysis to Capture Lexical Semantic Relation
3 . 学会等名 Proc. of EMNLP2018, pp.594-600 ( 国際学会 )
4 . 発表年 2018年

1 . 発表者名 Kazuya Ueki, Koji Hirakawa, Kotaro Kikuchi, and Tetsunori Kobayashi,
2 . 発表標題 Fine-grained Video Retrieval using Query Phrases: Waseda_Meisei TRECVID 2017 AVS System
3 . 学会等名 Proc. of ICPR2018 ( 国際学会 )
4 . 発表年 2018年

1 . 発表者名 George Awad, Alan Smeaton, Cees Snoek, Shin'ichi Satoh, Kazuya Ueki
2 . 発表標題 Video Recognition and Retrieval at the TRECVID Benchmark
3 . 学会等名 ECCV2018, Tutorial ( 招待講演 ) ( 国際学会 )
4 . 発表年 2018年

1. 発表者名 Yu Nakagome, Kazuya Ueki, Koji Hirakawa, Kotaro Kikuchi, Yoshihiko Hayashi, Tetsuji Ogawa, Tetsunori Kobayashi
2. 発表標題 Waseda_Meisei at TRECVID2018: Ad-hoc Video Search
3. 学会等名 Notebook paper of the TRECVID 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Kazuya Ueki
2. 発表標題 Latent Concept Extraction for Zero-shot Video Retrieval
3. 学会等名 Proc. of IVCNZ2018 (国際学会)
4. 発表年 2018年

1. 発表者名 黒澤 郁音, 菊池 康太郎, 小林 哲則, 林 良彦
2. 発表標題 言語情報を用いたCRFによるシーングラフ認識
3. 学会等名 第21回 画像の認識・理解シンポジウム (MIRU2018)
4. 発表年 2018年

1. 発表者名 平川 幸司, 菊池 康太郎, 植木 一也, 林 良彦, 小林 哲則
2. 発表標題 クエリ文からのコンセプトの選択に基づくアドホック動画検索
3. 学会等名 第21回 画像の認識・理解シンポジウム (MIRU2018)
4. 発表年 2018年

1. 発表者名 植木 一也
2. 発表標題 ゼロショット映像検索のための潜在的なコンセプトの抽出
3. 学会等名 第21回 画像の認識・理解シンポジウム (MIRU2018)
4. 発表年 2018年

1. 発表者名 長谷川美夏, 小林哲則, 林 良彦
2. 発表標題 語彙的含意関係識別による単語意味属性の推定
3. 学会等名 言語処理学会第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 黒澤郁音, 菊池康太郎, 小林哲則, 林 良彦
2. 発表標題 言語特徴量を利用したシーングラフ生成の効率的な計算機構
3. 学会等名 言語処理学会第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 平川幸司, 小林哲則, 林 良彦
2. 発表標題 動画からの動作キャプション生成における書き換え技術の適用
3. 学会等名 言語処理学会第25回年次大会 (NLP2019)
4. 発表年 2019年

1. 発表者名 植木 一也, 中込 優, 平川 幸司, 菊池 康太郎, 林 良彦, 小川 哲司, 小林 哲則
2. 発表標題 クエリ文によるゼロショット映像検索 - TRECVID 2018 AVSタスクの成果報告 -
3. 学会等名 動的画像処理実用化ワークショップ (DIA2019)
4. 発表年 2019年

1. 発表者名 Kentaro Kanada, Tetsunori Kobayashi and Yoshihiko Hayashi
2. 発表標題 Classifying Lexical-semantic Relationships by Exploiting Sense/Concept Representations
3. 学会等名 Workshop on Sense, Concept and Entity Representations and their Applications (SENSE) (April 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Mika Hasegawa, Tetsunori Kobayashi and Yoshihiko Hayashi
2. 発表標題 Incorporating visual features into word embeddings: A bimodal autoencoder-based approach
3. 学会等名 IWCS 2017 (September 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Kazuya Ueki, Koji Hirakawa, Kotaro Kikuchi, Tetsuji Ogawa and Tetsunori Kobayashi
2. 発表標題 Waseda_Meisei at TRECVID 2017: Ad-hoc Video Search
3. 学会等名 TRECVID 2017 (November 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Ikuto Kurosawa, Tatsunori Kobayashi and Yoshihiko Hayashi
2. 発表標題 A Neural Network Model for Detecting Inter-object Relationships
3. 学会等名 CVPR Language and Vision workshop (July 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 Kotaro Kikuchi, Naohiro Tawara, Tatsunori Kobayashi and Yoshihiko Hayashi
2. 発表標題 Word Vector Augmentation by its Definition for Zero-shot Image Classification
3. 学会等名 CVPR Language and Vision workshop (July 2017) (国際学会)
4. 発表年 2017年

1. 発表者名 金田健太郎, 小林哲則, 林 良彦
2. 発表標題 単語、語義、概念：意味タスクにおける分散表現の適用性
3. 学会等名 人工知能学会全国大会 (June 2017)
4. 発表年 2017年

1. 発表者名 菊池康太郎, 依直弘, 小林哲則
2. 発表標題 辞書定義文を用いたゼロショット一般物体認識
3. 学会等名 画像の認識・理解シンポジウム (MIRU) (August 2017)
4. 発表年 2017年

1. 発表者名 植木 一也, 平川 幸司, 菊池 康太郎, 小林 哲則
2. 発表標題 クエリ文を用いた詳細映像検索 -TRECVID 2017 AVSタスクの成果報告-
3. 学会等名 動的画像処理実用化ワークショップ(DIA2018) (March 2018)
4. 発表年 2018年

1. 発表者名 加藤恒昭
2. 発表標題 動詞語義の階層的分類に関する一考察
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

1. 発表者名 鷺尾 光樹, 加藤 恒昭
2. 発表標題 単語ペアと依存構造パスの共起モデリングを用いた語の意味関係の分類
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

1. 発表者名 金田健太郎, 小林哲則, 林 良彦
2. 発表標題 語義・概念の分散表現を利用したSemantic Taxonomy Enrichment
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

1. 発表者名 黒澤郁音, 菊池康太郎, 小林哲則, 林 良彦
2. 発表標題 共起性を利用した物体認識における言語情報の有効性
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

1. 発表者名 菊池康太郎, 林 良彦, 小林哲則
2. 発表標題 ゼロショット物体認識における辞書定義文の援用
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

1. 発表者名 平川幸司, 菊池康太郎, 植木一也, 林 良彦, 小林哲則
2. 発表標題 クエリ中の単語の語義絞り込みによる動画検索精度の向上
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

1. 発表者名 長谷川美夏, 小林哲則, 林 良彦
2. 発表標題 機械学習による単語概念の意味属性推定
3. 学会等名 第24回言語処理学会年次大会 (March 2018)
4. 発表年 2018年

〔図書〕 計3件

1. 著者名 Yoav Goldberg、加藤 恒昭、林 良彦、鷺尾 光樹、中林 明子	4. 発行年 2019年
2. 出版社 共立出版	5. 総ページ数 336
3. 書名 自然言語処理のための深層学習	

1. 著者名 van Erp, M., Hellmann, S., McCrae, J.P., Chiarcos, C., Choi, K.-S., Gracia, J., Hayashi, Y., Koide, S., Mendes, P., Paulheim, H., Takeda, H. (Eds.)	4. 発行年 2017年
2. 出版社 Knowledge Graphs and Language Technology	5. 総ページ数 152
3. 書名 Springer	

1. 著者名 Christopher D.Manning、Hinrich Schutze、加藤 恒昭、菊井 玄一郎、林 良彦、森 辰則	4. 発行年 2017年
2. 出版社 共立出版	5. 総ページ数 640
3. 書名 統計的自然言語処理の基礎	

〔産業財産権〕

〔その他〕

-

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	加藤 恒昭  (Kato Tsuneaki)  (60334299)	東京大学・大学院総合文化研究科・教授   (12601)	

6. 研究組織（つづき）

	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	小川 哲司  (Ogawa Tetsuji)  (70386598)	早稲田大学・理工学術院・教授    (32689)	
研究分担者	植木 一也  (Ueki Kazuya)  (80580638)	明星大学・情報学部・准教授    (32685)	

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