

令和 6 年 6 月 10 日現在

機関番号：82636

研究種目：若手研究

研究期間：2022～2023

課題番号：22K17978

研究課題名（和文）環境の階層型セマンティックモデリング

研究課題名（英文）Hierarchical Semantic Modeling of Environments

研究代表者

伊東 聖矢 (Ito, Seiya)

国立研究開発法人情報通信研究機構・ユニバーサルコミュニケーション研究所先進的リアリティ技術総合研究室
・研究員

研究者番号：30938709

交付決定額（研究期間全体）：（直接経費） 2,600,000円

研究成果の概要（和文）：本研究では、ロボットが周囲の環境を幾何的かつ意味的に理解し、柔軟な判断を行うための基盤技術として、概念階層に基づく画像認識技術を開発した。この技術により、異なるレベルの意味情報をもつ3次元環境地図の構築が可能となる。また、セマンティックモデリングの性能を評価するために、静的な環境に実世界の動的要素を組み込んだ新しいベンチマークデータセットを構築し、動的な環境における分析を可能にした。さらに、意味情報の増加に伴い環境地図データが肥大化する問題に対処するために、物体の幾何情報と意味情報を効率的に圧縮する技術を開発した。

研究成果の学術的意義や社会的意義

環境のセマンティックモデリングに関する研究はコンピュータビジョンやロボティクス分野で進められてきているが、意味推定には既存の画像認識技術を利用しており、その多くは認識対象とするクラスを同列に扱っている。本研究では、認識対象とするクラスに基づいて、抽象度の高いクラスから具体的なクラスまで、異なる概念レベルでの意味理解を可能にした。これにより、ロボットが複雑な環境や多様な概念を効率的に処理できるようになるため、学術的意義は大きいといえる。また、ロボットの複雑な環境に対する柔軟な理解および適応能力を向上させ、ロボットの実用化に貢献することが期待される。

研究成果の概要（英文）：This study developed a foundational technology for robots to understand their surrounding environments geometrically and semantically, facilitating flexible decision making. The core of this technology is an image recognition algorithm based on conceptual hierarchies, which allows the construction of 3D environment maps that encompass various levels of semantic information. To evaluate the performance of semantic modeling, a new benchmark dataset was constructed, which combines real-world dynamic elements and static environments, enabling analysis in dynamic contexts. Furthermore, to address the issue of data bloat caused by increasing semantic information, this study developed a method for efficiently compressing both the geometric and semantic data of objects.

研究分野：コンピュータビジョン

キーワード：3Dモデリング 意味情報 概念階層

※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

少子高齢化に伴う労働力不足は社会課題の一つであり、その対策として、高度な人工知能を備え、柔軟に判断して行動することができるロボットの活躍が期待されている。ロボットが自律的に行動するためには、ロボットが行動する空間の環境地図が不可欠である。しかし、ロボットが常に最新の環境地図を保持しているとは限らない。したがって、ロボットは搭載しているセンサを使用して、その場で環境を幾何学的および意味的に理解する能力が求められる。

近年、コンピュータビジョンやロボティクス分野において、画像から意味情報をもつ 3 次元環境地図を構築するセマンティックモデリングの開発が進展している。このセマンティックモデリング技術により、ロボットは画像から周辺にある物体や状況を認識し、それに基づいて適応的に行動を決定することが可能になる。現在のセマンティックモデリングで使用されている画像認識技術の多くは、認識対象となる意味クラスを事前に定義し、すべての意味クラスが同列に扱われている。しかしながら、実応用においては、目的や状況に応じて必要な意味クラスの粒度が異なる。ここで、意味クラスの粒度とは、ある意味クラスに関する概念の詳細さを指す。例えば、「椅子」を認識する場合、状況によって、「家具」のような抽象度の高いクラスとして認識すれば十分な場合もあれば、「肘掛け椅子」のような詳細なクラスとして認識することが必要な場合もある。このような状況に応じて必要な意味の粒度で認識する技術は確立されておらず、これまで意味クラスの定義は人の手で慎重に設計されていた。

また、セマンティックモデリングにおいて、環境地図に埋め込む情報は多様化してきている。意味情報を含む環境地図データは、扱う意味情報の種類の増加に伴って肥大化するため、3 次元形状と意味情報を効率よく表現することが求められている。

2. 研究の目的

本研究の目的は、ロボットが高度な状況判断を行うために、形状情報と意味情報を集約した環境地図を構築することである。特に、本研究では、意味情報を概念階層に基づいて階層的に理解することが可能な「階層型セマンティックモデリング」の実現を目指す。従来のセマンティックモデリングにおける意味理解は、事前に規定された意味クラスに基づいて行われていた。それに対し、本研究では、概念階層を活用して、規定された意味クラスに基づき、抽象的なクラス（上位概念）から具体的なクラス（下位概念）までをすべて認識対象とする認識技術を開発し、ロボットが柔軟な判断を下せる枠組みを構築することを目指している。

このように、規定された意味クラスだけでなく、その上位概念の意味クラスを認識対象とすることで、従来よりも多くの意味クラスを扱うことになり、その結果、環境地図データは肥大化する。そこで、本研究では、環境に存在する物体と意味情報を低次元の特徴ベクトルで表現する方法を開発する。環境地図データを物体単位で分割し、各物体を冗長性を排除した低次元の特徴ベクトルで表現することで、データ量の削減が期待できる。

3. 研究の方法

階層型セマンティックモデリングを実現するためには、概念階層に基づいて意味情報を理解する技術を確立する必要がある。本研究では、あらかじめ定義された認識対象クラスから、概念階層を自動構築し、その概念階層に基づいて画像認識を行う技術を開発する。従来の画像認識と比較して、この技術がどの程度概念階層に従って認識できているかを、様々なデータセットを用いて検証する。

また、これまでのセマンティックモデリングは主に静的なシーンのみで評価されており、動物体が存在する環境における評価が十分に行えていなかった。そこで、静的シーンに動物体を合成する手法を開発し、実世界を模した評価用データセットを構築する。このデータセットを用いることで、動物体がセマンティックモデリングに与える影響を定量的に分析することが可能になる。

さらに、物体の 3 次元モデルを低次元の特徴ベクトルで表現し、その特徴ベクトルから物体の詳細な意味情報を解析する手法を開発する。具体的には、ニューラルネットワークを用いて特徴ベクトルを分割し、分割した特徴ベクトルが物体の部分形状、すなわちパーツとなるように教師なし学習を行う。そして、どの程度パーツを再現できるかを既存のパーツ情報を含むデータセットを用いて評価する。

4. 研究成果

本研究の主の成果として、次の 3 つを開発した。

- (1) 複雑な概念階層を扱うことが可能な階層画像分類法
- (2) セマンティックモデリングの評価のための実世界シーンを模した動的シーン生成手法
- (3) 階層的パーツ分割可能な低次元ベクトルによる物体形状の表現方法

これらの成果について以下で説明する。

(1) 複雑な概念階層を扱うことが可能な階層画像分類法

階層画像分類は、画像の理解を促進するための補助情報として階層構造（主に階層木）を利用して画像进行分类する問題である。従来の階層画像分類法では、画像分類モデルの構造に適合するように階層木を手動で調整する必要があった。例えば、マルチスケールの特徴抽出器を用いた画像分類モデルでは、スケールと階層の深さが対応するように階層木が構築されていた。そのため、階層木には深さを調整するために導入された不合理なノードが含まれるという問題があった。

この問題に対し、本研究では、概念階層木の各ノードに対する予測確率を求め、それらを概念階層に沿って統合することで最終的なクラスを予測する新しい方法を提案した（図 1）。提案手法では、画像分類モデルに合わせて階層木の深さを調整する必要がなくなり、従来よりも幅広い概念階層を使用することが可能になる。そこで、本研究ではさらに、語彙データベースを用いて概念階層を自動構築する手法を提案した。この自動階層構築法は、予測対象となる詳細クラスに基づきクラス間に共通する上位概念を網羅するように階層を構築する初期化と、構築した階層から不合理なノードを取り除く圧縮からなる。図 2(a)は手動で設計された概念階層、図 2(b)は提案手法によって構築された概念階層である。手動で構築した概念階層は、階層木の深さが統一されているものの、“birds”の下位概念が“bird”となっており、階層の深さを調整するための不合理なノードが含まれている。一方で、提案手法で自動構築した概念階層は、階層木の深さは統一されていないが、語彙データベースに基づいているため、妥当な階層を得られる。

提案した画像認識モデルについて、手動で構築した概念階層と自動で構築した概念階層を用いて比較した。その結果、自動構築した概念階層を用いた場合の精度が、手動構築した概念階層を用いた場合の精度と同等以上であることを確認した。また、従来の階層画像分類モデルと比較し、提案手法が最高精度であることを確認した。さらに、詳細に分析した結果、異なる概念レベルを網羅するように予測できていることが示された。

この手法では、階層クラス分類と最終的なラベルの決定を別々の処理として行う必要がある。そこで、階層画像分類の問題設定を再考し、多クラスランキング問題とみなすことで、階層画像分類とラベル決定を同時に学習する手法を開発した。これにより、異なる概念レベルで意味を予測できるだけでなく、その中から適切なラベルを選択することができる、階層型セマンティックモデリングの基盤技術を確立した。

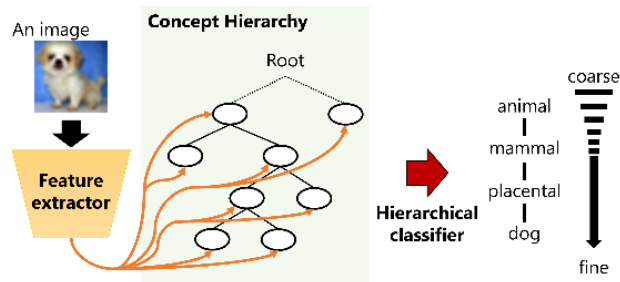
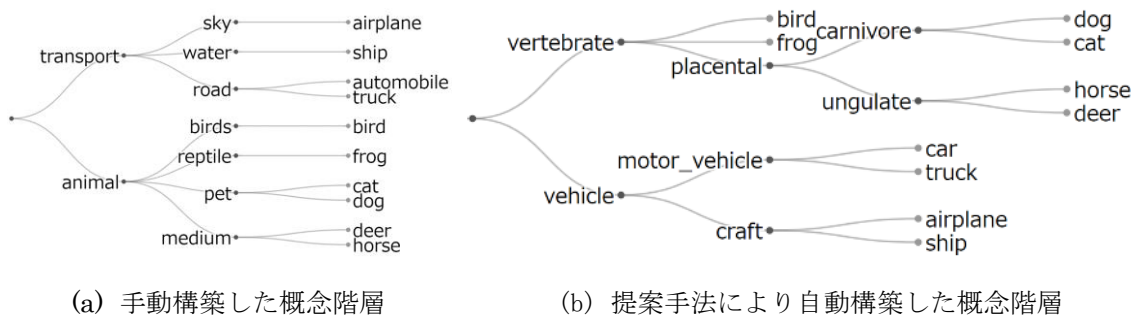


図 1 概念階層に基づく階層画像分類モデルの概要



(a) 手動構築した概念階層

(b) 提案手法により自動構築した概念階層

図 2 概念階層

(2) セマンティックモデリングの評価のための実世界シーンを模した動的シーン生成手法

環境のセマンティックモデリングは、コンピュータビジョンやロボティクス分野で広く研究されているが、その評価はこれまで主に静的シーンで行われてきた。しかし、実世界シーンには静的な物体だけでなく、時間とともに変化する動的要素も多く含まれている。実世界シーンは非常に複雑であり、多種多様な動的要素を網羅したデータセットを構築することは困難である。また、動的シーンに含まれるすべての要素を正確にアノテーションすることも難しい。さらに、実世界でのデータ収集においてはプライバシーや倫理的問題を考慮する必要がある。

例えば、人々の動きや行動を記録する場合は、プライバシー保護のためにデータを匿名化しなければならない。

このような課題に対し、本研究では、静的シーンに動的な物体を組み込むことで動的シーンを生成する手法を開発した。具体的には、人物を動的な物体として扱い、生成モデルを利用して人物の様々な動きを生成する。図3は生成された人物の動作の例であり、生成モデルが環境に応じて行動する人物を生成していることがわかる。ここで、静的シーンには既存のデータセットをそのまま活用することができる。また、動的な物体は動的シーンと静的シーンの差分を計算することでアノテーションを付与することができる(図4)。この手法により、動的シーンの生成を効率的に行うことができ、セマンティックモデリングの評価において、より現実的な評価基準が得られる。

本研究では、特に動的な物体がセマンティックモデリングに与える影響に着目し、画像中の動物体の割合に対するセマンティックモデリングの性能を分析した。図5は左上の領域で人物が動くシーンで構築したモデルの例であり、静的シーンの結果と比べて誤った意味ラベルが付与されていることが確認できる。動的物体がセマンティックモデリングの性能に与える影響は、これまでは定性的な分析に留まっていたが、本研究により、全体の性能への影響だけでなく、各意味ラベルへの影響を定量的に分析することが可能になった。



図3 生成モデルで生成した人物の動作



図4 動的物体のアノテーション付与

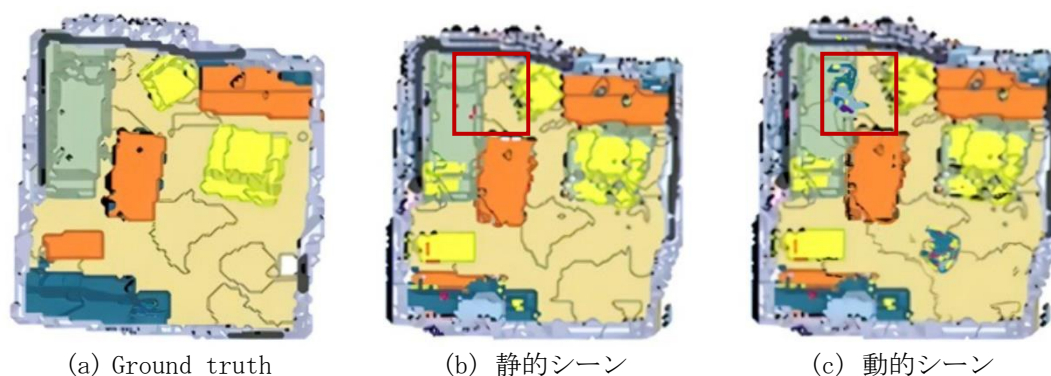


図5 各シーンでのセマンティックモデリング結果

(3) 階層的パーツ分割可能な低次元ベクトルによる物体形状の表現方法

環境地図は自律移動ロボットや自動運転における周辺の環境の把握に重要である。しかし、大規模なシーンや詳細な意味情報を含む環境地図を構築すると、データ量が膨大になる。このような課題を解決するために、必要な情報を維持しつつデータ量を削減する圧縮技術の開発が求められている。有効な方法の一つは、環境を一つのモデルとして表現するのではなく、環境

内の物体を個々に分割し、それぞれを低次元のベクトルとして表現することである。これにより、個々の物体の重要な情報を抽出し、不要な情報を除去することができるため、全体のデータ量を軽量化することができる。

このような背景から、本研究では、物体の3次元モデルを低次元なベクトルで表現し、そのベクトルを用いて物体を階層的に分割する手法を開発した（図 6）。提案手法では、まず3次元モデルから低次元の特徴ベクトルを抽出し、その特徴ベクトルを2つの特徴ベクトルに再帰的に分割していく。最初に得られる特徴ベクトルは物体全体の3次元形状を表し、分割された特徴ベクトルがそれぞれパーツを表すように学習する。さらに、提案手法は、3次元形状の最小構成要素であるパーツの度合いを予測するパーツプリミティブ度合いネットワークを導入し、分割された特徴ベクトルから復元される3次元形状がパーツとなることを促進する。このパーツプリミティブ度合いネットワークは合成データセットを用いて学習される。また、このネットワークは物体が過剰に分割されているかどうかの判定にも使用できる。

提案手法をパーツラベルをもつ既存のデータセットで評価した結果、パーツに関する教師ラベルを用いていないにもかかわらず、意味的なパーツ分割が可能であることを確認した（図 7）。また、過分割と判定された形状を統合することで、精度が向上することを示した。

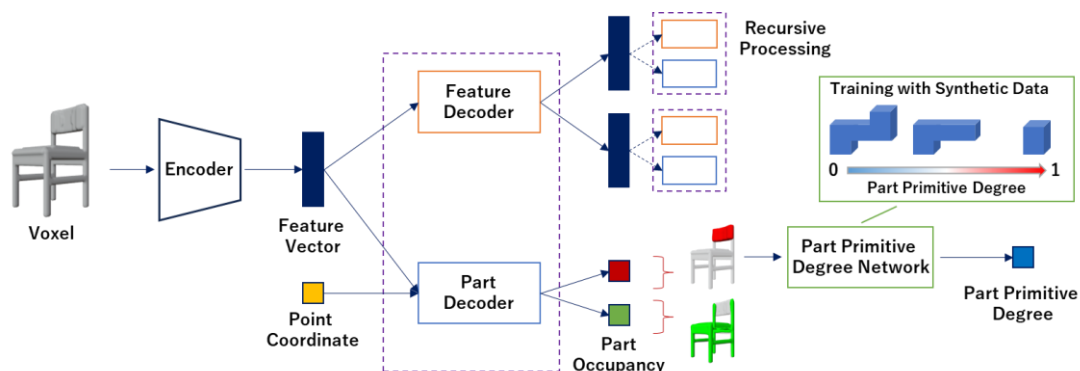


図 6 物体の3次元形状を低次元特徴ベクトルで表現する手法の概要

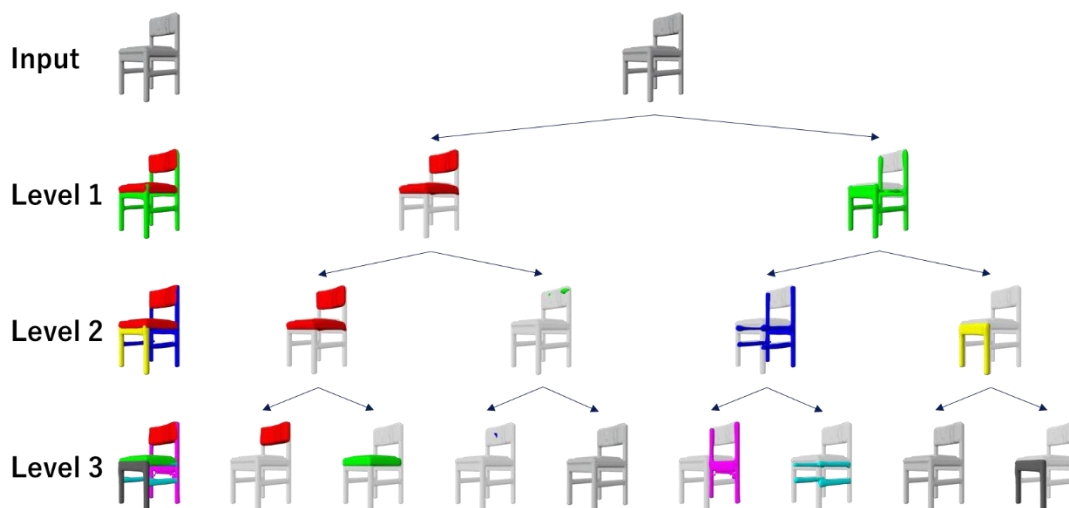


図 7 階層的なパーツ分割

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計13件（うち招待講演 1件／うち国際学会 6件）

1. 発表者名 Mami Kikuchi, Seiya Ito, Naoshi Kaneko, Kazuhiko Sumi
2. 発表標題 Part-Level Single-View 3D Shape Reconstruction with Multiple Types of Primitives
3. 学会等名 Sixteenth International Conference on Quality Control by Artificial Vision (QCAV 2023) (国際学会)
4. 発表年 2023年

1. 発表者名 山崎 禎晃, 伊東 聖矢, 大原 剛三
2. 発表標題 概念辞書から自動構築した概念階層に基づく画像分類
3. 学会等名 第29回画像センシングシンポジウム (SSII2023)
4. 発表年 2023年

1. 発表者名 伊東 聖矢
2. 発表標題 画像を用いた環境のセマンティックモデリング
3. 学会等名 精密工学会 画像応用技術専門委員会 サマーセミナー2023 (招待講演)
4. 発表年 2023年

1. 発表者名 周 沁園, 伊東 聖矢, 鷺見 和彦
2. 発表標題 動物体がSemantic Scene Completionに与える影響の定量化手法の検討
3. 学会等名 精密工学会 画像応用技術専門委員会 サマーセミナー2023
4. 発表年 2023年

1. 発表者名 Qin-yuan Zhou, Seiya Ito, Kazuhiko Sumi
2. 発表標題 A Benchmark for 3D Reconstruction with Semantic Completion in Dynamic Environments
3. 学会等名 30th International Workshop on Frontiers of Computer Vision (IW-FCV 2024) (国際学会)
4. 発表年 2024年

1. 発表者名 Yusho Miyashita, Seiya Ito, Kazuhiko Sumi
2. 発表標題 Speeding Up Semantic Scene Completion: A Completion by Region Proposal Approach
3. 学会等名 30th International Workshop on Frontiers of Computer Vision (IW-FCV 2024) (国際学会)
4. 発表年 2024年

1. 発表者名 Mami Kikuchi, Seiya Ito, Naoshi Kaneko, Kazuhiko Sumi
2. 発表標題 Learning Part Primitive Degrees for Hierarchical 3D Object Partitioning
3. 学会等名 30th International Workshop on Frontiers of Computer Vision (IW-FCV 2024) (国際学会)
4. 発表年 2024年

1. 発表者名 Tomoaki Yamazaki, Seiya Ito, Kouzou Ohara
2. 発表標題 Tailoring Ranking-based Losses to Capture Superior-Subordinate Relationships for Hierarchical Image Classification
3. 学会等名 30th International Workshop on Frontiers of Computer Vision (IW-FCV 2024) (国際学会)
4. 発表年 2024年

1. 発表者名 菊地真美, 伊東聖矢, 金子直史, 鷺見和彦
2. 発表標題 複数種類のプリミティブによるパーツレベルの単視点3次元形状再構成
3. 学会等名 第21回情報科学技術フォーラム (FIT2022)
4. 発表年 2022年

1. 発表者名 Tomoaki Yamazaki, Seiya Ito, Kouzou Ohara
2. 発表標題 Hierarchical Image Classification with Conceptual Hierarchies Generated via Lexical Databases
3. 学会等名 29th International Workshop on Frontiers of Computer Vision (IW-FCV 2023) (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Dyna3DBench https://github.com/seiyaito/Dyna3DBench
--

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

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
---------	---------