

SUPS: 自律走行のための地下駐車場シナリオデータセットのシミュレーション

Jiawei Hou¹, Qi Chen¹, Yurong Cheng¹, Guang Chen², Xiangyang Xue³, Taiping Zeng³, Jian Pu^{3*}

¹ School of Computer Science, Fudan University, Shanghai, China.

² FAW (Nanjing) Technology Development Co.,Ltd, Nanjing, China.

³ Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China.

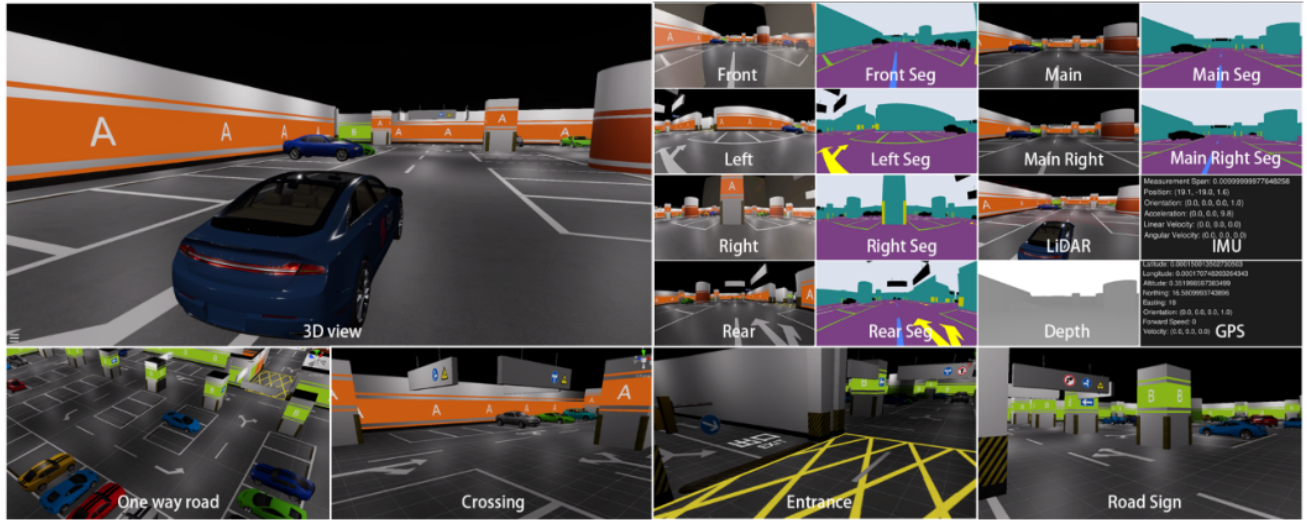


図1. センサの完全なセットからの出力の例。左上隅はシミュレーションシーンにおける車両の3Dビューである。右上には4列の画像があり、左から右へ、1列目は4台の魚眼カメラで撮影されたフレーム、2列目は前列の対応するセマンティックカメラで撮影されたセマンティックセグメンテーションのグラントゥールース、3列目は前方ステレオピンホールカメラ、LiDARスキャン、深度カメラからのフレーム、最後の列はステレオカメラのセグメンテーションのグラントゥールースとIMUとGPSからの測定値を示している。下の4枚の写真は、一方方向道路、横断歩道、入り口、道路標識など、仮想シーンにおけるいくつかの繊細なシナリオを示している。

概要- 自動地下駐車は、自律走行の範囲が拡大するにつれて、大きな注目を集めている。自動車両は、環境情報を取得し、その位置を追跡し、シナリオの信頼できるマップを構築することになっている。主流のソリューションは、よく訓練されたニューラルネットワークと同時定位マッピング(SLAM)手法で構成されており、多数の注意深くラベル付けされた画像と複数のセンサー推定を必要とする。しかし、SLAMタスクとセマンティックセグメンテーションや駐車場検出などの知覚タスクの両方をサポートする、複数のセンサーとラベル付けされた画像を持つ地下駐車場シナリオデータセットが不足している。本論文では、複数のセンサと複数の意味ラベルを持つ複数のタスクを、タイムスタンプに従って連続した画像と整理させた、地下自動駐車のためのシミュレーションデータセットであるSUPSを紹介する。我々は、環境の多様性と仮想シーンにおけるセンサーの多様性とアクセシビリティを持つ既存のデータセットの欠陥をカバーすることを意図している。具体的には、周囲の魚眼カメラ4台、前方ピンホールカメラ2台、深度カメラ1台、LiDAR、慣性計測ユニット(IMU)、GNSSのデータからフレームを記録する。オブジェクト、特に矢印、駐車線、車線、スピードバンプなどの地面標識に対して、ピクセルレベルの意味ラベルが提供される。知覚、3D再構成、奥行き推定、SLAM、その他の相対タスクは、我々のデータセットによってサポートされている。

また、我々のデータセットで、最先端のSLAMアルゴリズムと知覚モデルを評価する。最後に、Unity Engineをベースに構築した仮想3Dシーンをオープンソース化し、データセットを<https://github.com/jarvishou829/SUPS>で公開する。

I. INTRODUCTION

自動地下駐車は、自律走行に不可欠な部分の一つである。このプロセスにおいて、地下駐車シナリオから環境情報を取得し、センサーからの入力を処理し、計画を立てるために、自家用車両を提案する。この完全な手順に対処するためには、セマンティックセグメンテーション、駐車場検出、同時定位マッピング(SLAM)など、様々なアルゴリズムが必要である。しかし、これらのアルゴリズムを適用することは、地下駐車場シナリオでは困難である可能性がある。このようなシナリオでは、障害物のある狭い道路、まばゆいや薄暗い照明のある場所、質感の悪い壁などがセンサー測定のノイズや不確かさを増大させる。

TABLE I

S様々な自律走行データセットと構成およびタスクの概要

Task/Info	Quantity	PS2.0 [1]	Cityscapes [2]	TUM VI [3]	KITTI [4]	WoodScape [5]	VIODE [6]	SUPS(Ours)
Capture Information	Year	2018	2016	2018	2012/14/15	2018/19	2021	2021/22
	Carrier	Car	Car	UAV	Car	Car	UAV*	Car
	Other sensors	-	-	IMU	1 LiDAR GPS	1 LiDAR GNSS IMU	IMU* Segmentation	1 LiDAR* GPS*,IMU* Segmentation*
	Ground-truth Trajectory	-	-	MoCap (partial)	Fused IMU, GNSS	Fused IMU, GNSS	Simulation	Simulation
Camera Information	Cameras	-	2	2	4	4	1	6
	Type	Surround view	-	Stereo	Stereo	Surround view	Stereo*	Stereo* Surround view*
Parking Slot Detection	Support	✓	-	-	-	-	-	✓
Semantic Segmentation	Support	-	✓	-	✓	✓	✓	✓
Depth Estimation	Support	-	-	✓	✓	✓	-	✓
Visual SLAM	Support	-	-	✓	✓	△	✓	✓
LOAM SLAM	Support	-	-	-	-	△	-	✓

4我々の知る限り、SLAMタスクのためのいくつかのシーケンスは、WoodScape [5]ではまだオープンソースではありません * シミュレーションされたキャリアとセンサー。

[7]. 環境の構造と内容は、ランドマークがある都市の街並みに比べて繰り返しが多く、特徴量の照合や場所の認識が困難である。さらに、全地球航法衛星システム (GNSS) の信号遮蔽、薄暗い条件、地盤の反射 [8] [9] により、車両は位置の信頼性の低い推定結果を得る可能性がある。

最近、研究者は、ディープラーニングを応用して、車両が周囲のシナリオを認識できるようにすることで、これらの課題を部分的に解決できることを発見した [10] [11] [12]。具体的には、意味情報は自律走行車の知覚能力を向上させ、その見返りとして、正確な姿勢推定も知覚に役立つ可能性がある。しかし、一般的に使用されるデータセットの多くは、SLAMタスクではタイムスタンプ付きの連続画像、視覚認識タスクでは離散画像を提供しており、ユーザが単一のデータセットを用いて機能的な自動駐車システムを開発することは困難である。また、地下駐車シナリオでは、壁、柱、走行可能エリア、レーン、駐車場、矢印、段差など、障害物や地上標識の意味情報が密集している。これらの項目を完全にラベル付けした既存のデータセットはわずかであり、人間の労力と時間を大量に必要とする。

SLAMと知覚タスクの両方をサポートするデータセットの不足をカバーするために、本論文では、地下自動駐車システムの開発における問題に対処するために、複数のセンサーと複数の意味ラベルを通して複数のタスクをサポートする新しいデータセットであるSUPSを紹介する。SUPSで使用される仮想シーンは、LG SVL [13] を介してシミュレートされる。LG SVLは、車両とセンサーの非常に柔軟な構成を持つ自律走行車シミュレーションプラットフォームである。シミュレーションプラットフォームは、環境変動だけでなく、仮想センサーの多様性とアクセシビリティという利点がある。私たちが構築した仮想地下駐車シナリオは、照明、質感、内容、シナリオの複雑さにおいて、現実の地下駐車シナリオと類似している。

魚眼カメラ、ピンホールカメラ、深度カメラ、LiDAR、慣性計測ユニット (IMU)、GNSS などの複数のセンサーが記録間で起動される。このデータセットでは、ピクセルレベルのセマンティックセグメンテーション、駐車場検出、奥行き推定、ビジュアルSLAM、LiDARベースのSLAMタスクがサポートされている。

本研究の主な貢献は以下のようにまとめられる：

- 我々は、複数のセンサーと複数の意味ラベルを持つ複数のタスクを、タイムスタンプに従って連続した画像に整列させた、地下自動駐車のためのシミュレーションデータセットであるSUPSを発表する。地下駐車シナリオにおけるSLAMと知覚アルゴリズムの頑健性と精度のベンチマークが可能である。
- 我々は、いくつかの最先端のセマンティックセグメンテーションとSLAMアルゴリズムを我々のデータセットで評価し、その実用性と地下駐車シナリオにおける困難な困難性を示す。
- SUPSデータセットとシミュレーションの地下駐車場シナリオ全体をオープンソース化し、研究者が特定のタスクのために独自に設計した変更を行うことができるようにした。

II. RELATED WORK

本節では、まず自律走行タスクの主流ベンチマークデータセットを簡単にレビューし、次に自律走行システムにおける周辺知覚システム、意味情報、SLAMアルゴリズムの統合について議論する。

A. 自律走行データセット

既存の自律走行データセットは、キャリア、環境、センサーなど多様である。しかし、ほとんどのデータセットはSLAMか知覚タスクのどちらか一方のみを対象として設計されている。例えば、EuRoC MAV [14]、TUM VI [3]、nuScenes [15]は、土地標識のセマンティックセグメンテーションタスクのためのラベル情報を提供していない。都市景観 [2]、nuScenes [15]、

ApolloScape [16]、BDD100k [17]は深度推定とSLAMタスクをサポートしていない。人件費と資源費が極めて高いため、複数のタスクを組み合わせる研究している研究者はほとんどいない。

地下駐車シナリオを含むデータセットはわずかであった。VIODE [6]は地下シーンを提供したが、データセットは無人航空機(UAV)で記録されており、いくつかの自律走行アルゴリズムには適していない。PS2.0[1]は駐車場検出タスクのデータセットであり、地下駐車場シナリオを含むが、SLAMタスクはサポートされていない。このように、地下自動駐車には明らかな欠陥がある。同時に、KITTI [4]、Cityscapes [2]、BDD100k [17]などの主流の自律走行データセットでは、サラウンドビューカメラは適用されていない。

要約すると、オープンソースの自律走行データセットは、地下駐車シナリオにおいて、SLAMと知覚タスクに同時に不十分であった。仮想シーンを活用し、我々のデータセットは、実際の地下駐車シナリオがどのようなものかを考慮し、多様なタスクに複数のセンサーを提供することで、この問題に対する比較的完全な解決策として機能する。Tab. Iは、既存のデータセットに対する我々のデータセットの新規性を示している。

B. 自律走行のための知覚システム

360°の知覚を持つ複雑なアプリケーションを扱うことへの関心が高まっている。GNN[18]、[19]は、鳥瞰(BEV)[10]の視界における駐車スロット検出を改善するために適用されている。マルチピンホールカメラは、モーションプランニングタスクを解決するためのエンドツーエンドのアルゴリズムに使用されています[11]。同時に、セマンティックセグメンテーション、検出、シーンエンコーディングなどの知覚アルゴリズムは、SLAMタスクにおいて激しい議論を引き起こしている。Mask-SLAMは、セマンティックセグメンテーション[20]を用いて、情報の乱れを避けるために、標準的なアーキテクチャにマスクを結合する。DynaSLAMは、セマンティックセグメンテーション[21]を用いて、動的なオブジェクトを除去した。VIODEは、いくつかの移動物体の特徴点の使用を排除するVINS-Maskを提案した[6]。しかし、我々の知る限り、特徴抽出やフレームマッチングにおいて意味情報を利用するオープンソースのアルゴリズムは実に少ない。このようなアルゴリズム開発の障害の一つは、トラッキングとマッピングのための豊富なラベル付き画像を持つデータセットがないことである。さらに、既存のデータセットは、主に走行可能領域、車線、動的な物体(他の車両、自転車、人など)のラベルに焦点を当てている。これらの特徴は、位置を追跡し、使用可能なマップを構築するのにほとんど役立たない。

III. データセットの概要

本節では、4つのセクションからなるデータセットを詳細に紹介する。まず、このデータセットを構築する動機について述べる。次に、プラットフォーム、シミュレータ、キャリア設定、データ処理方法によるデータセット生成手順を紹介する。第三に、データ列を記録し、集合を分割する戦略について説明する。最後に、データセットの全内容を詳しく説明する。



図2. シミュレーションシーンの概要領域Aは右側に示されたオレンジ色の部分であり、領域Bは左側に示された緑色の部分である。

A. Motivation

1) 周囲の知覚: 駐車タスクを完了するために、自動運転車はブラインドコーナーなしで周囲の環境の情報を取得しなければならない。360°知覚システムは、車両を取り囲む4台の魚眼カメラ、前方に設置されたステレオピンホールカメラ、上方に設置された1台のLiDARで構成されている。ピッチ角が下向きの魚眼カメラは、FOVが大きく、地面の標識をよりよく認識するために使用される。

2) マルチタスクのためのマルチセンサー: SLAMタスクと様々な視覚タスクは、自律駐車の問題において、どちらも困難であり、互いに不可分である。セマンティックSLAMとマルチモーダルセンシングは、視覚、深度、測距、速度測定をカバーするセンサーの完全なセットと同様に、多くのよくラベル付けされたフレームを必要とする。仮想シーンにおけるシミュレーション・マルチセンサー・システムは、GPS、IMU、LiDAR、深度測定、およびタイムスタンプに従って連続する画像ごとに整列されたグラントゥルースのセマンティック・セグメンテーションを提供する。センサーの詳細な構成は、サブセクションIII-Bで紹介する。

3) 変更可能なシミュレーションシーン: 我々のシミュレートした地下駐車シナリオは、Unity Engine [22]に基づいて構築された。ユーザはシーンを簡単に変更し、特定のアプリケーションのために再構築することができる。例えば、駐車場の構造、交通標識、照明、意味ラベル、移動ルートはすべて変更可能である。既存のシーンの概要については、サブセクションIII-Bで説明する。

B. Dataset generation

SUPSデータセット生成手順には、仮想シーン構築、シミュレータ設定、記録と処理の3つのステップが含まれる。これらのステップは、Ubuntu 18.04、Intel(R) Core i7-11800H CPU、GeForce RTX 3060 (Laptop) GPUを搭載したコンピュータで完了した。

1) 仮想シーンの構築: Unity Engine [22]を用いて、地下駐車シナリオの3Dシミュレーションシーンを構築した。

TABLE II
Dエゴ車両モデルにおけるセンサーの詳細

Sensor name	Position	Sampling Frequency
Fisheye camera 1	front	20 Hz
Fisheye camera 2	left	20 Hz
Fisheye camera 3	right	20 Hz
Fisheye camera 4	rear	20 Hz
Pinhole camera 1	top	20 Hz
Pinhole camera 2	top-right	20 Hz
Depth camera	top	20 Hz
GPSオドメトリセンサー	rear axle center	30 Hz
IMU	rear axle center	100 Hz
LiDAR	top-behind	7 Hz(rotation)

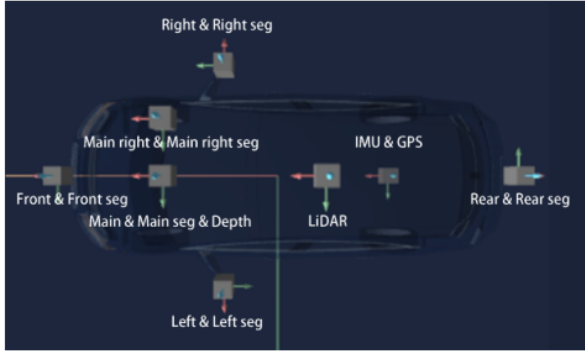


図3. エゴ・ビークル上のセンサーの位置カメラの名前は略して書かれている。例えば、Front & Front seg は、Front カメラと Front セマンティックセグメンテーションカメラを意味する。

図2に示すように、シーン全体はエリアAとエリアBの2つから構成される。エリアAは2635m²を占め、道路長は120mである。Bエリアは2015m²、道路は195m。シーン全体の走行可能な道路は約350mで、駐車場の面積は4650m²である。SLAMタスクにおいて、ループ閉鎖制約を与えるためのいくつかの異なる道路ループがシーンに存在する。シーンを豊かにするために、壁、柱、静止車両、駐車場、車線、走行可能エリア、衝突回避ストリップ、スピードバンプ、矢印、道路標識などの項目を含む。さらに、シミュレートされたシーンにおける知覚とSLAMタスクには、横断、入口、一方方向道路、ダズル、薄暗いエリア、低テクスチャエリアなど、現実世界の地下駐車シナリオに似た困難なシナリオがある(図1参照)。さらに、実際の地下駐車シナリオを最大限に回復するために、天井にスポットライトを当てる。

2) シミュレータのセットアップ LGSVL[13]プラットフォーム上で、Lincoln2017NKZ自律走行車モデルを作成し、複数のセンサーを適用する(センサーの詳細は表IIに示す)。エゴ・カーのセンサーの概要を図3に示す。また、センサーの固有パラメータと外在パラメータをすべて提供する。セマンティックセグメンテーションタスクのグラントゥールースを得るために、対応する4つの魚眼カメラと2つのピンホールカメラと同じパラメータを共有するセマンティックセグメンテーションカメラを自我モデル上に6台追加設置する。図1は、全センサーからの観測と測定の例である。



図4. 駐車場検出データフォームの例。左はオリジナル画像、内側はグラントゥールース、右はGCN[10]のフォーム説明である。

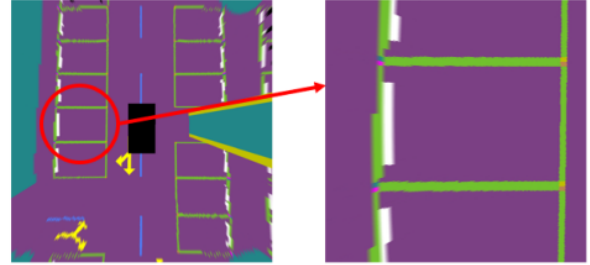


図5. シングルパーキングスロットの説明。駐車場の4つの角には、駐車場の幅と角の順番を示す異なる色の四角が付けられている。中心線、オブジェクトレベルのマスク、ベクトルを持つコーナーは、これらのラベルで簡単に生成できる。

3) データ処理: 既存のデータセットのうち、駐車枠検出に十分な情報を提供するものはわずかである。現実の世界では、駐車場が車や障害物に覆われたり、隠されたりすることが一般的であり、人間の労働者がそれを正確にレッテルを貼ることは困難である。ほとんどのデータセットにおいて、BEVのFOVは道路脇の完全な駐車スペースをカバーできない。同時に、アルゴリズムによって駐車スペースの記述は様々であった。例えば、GCN-parking-slotは、入口側に2つのコーナーがあり、コーナーの対応がある駐車場空間を記述している[10]。我々のデータセットは、駐車場を分離するためのすべての情報を含む、駐車場の適応可能な説明を提供する(図5参照)。ベクトル、中心線、オブジェクトレベルのマスクを持つコーナーは、順序付けられた点座標で簡単に抽出できる。例えば、GCN[10]の記述形式を用いた駐車場スロット検出のためのデータを提供します(詳細は図4参照)。さらに、鳥瞰(BEV)画像は、元のカメラフレーム以外に、自動車を操作する際の知覚や意思決定に特に役立つため、逆遠近マッピング(IPM)法によって投影されたBEV画像とセグメンテーションのグラントゥールース(図7参照)を提供する。

C. 取得戦略

1) SLAMタスク: 2つの速度レベルにおける3つのルートがデータセットに記録されている。図6に3つの走行経路(ループA、ループB、ループC)の詳細を示す。ループAは長さ約160mのルートで1つのループクロージャーを持ち、ループBは長さ約250mのルートで2つのループ(長さ60mと長さ150m)を持つ。ループCにはいくつかの複雑なループがあり、図6に示すようにルートは約700mである。



(a) Loop A and Loop B

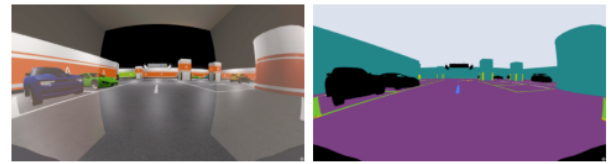


(b) Loop C

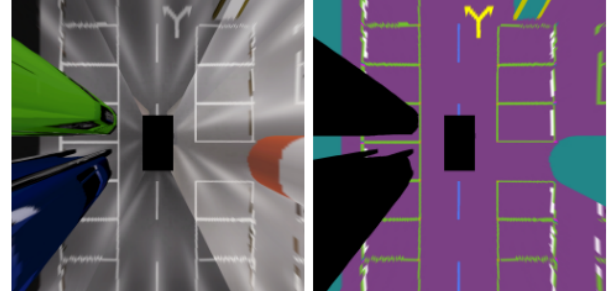
図6. 3つのルートと走行シーケンスの概要。緑色の点は自我がスタートする場所、赤色の点は自我が停止する場所である。交差点の各白点は、エゴ・カーが通過する経路を説明するために、AからMまでの大文字でラベル付けされている。(a)が示すように、ループAでは、エゴ・カーはスタート→C→D→H→G→エンドを走行する。ループBでは、エゴ・カーはスタート→B→A→E→F→J→I→K→L→エンドを走行する。(b)が示すように、ループCにおいて、自転車はスタート→C→G→H→D→B→L→M→C→B→F→E→A→B→J→I→K→M→Endを走行する。

ほとんどの地下駐車シナリオの速度制限は5km/hであるため、記録データの高速レベルは5km/hであり、車は低速レベルで3.5km/hで走行する。さらに、いくつかのSLAMアルゴリズム (Vins-MONO [23]など)は、初期化時に十分なIMU励起を必要とすることを考慮し、エゴ・カーがルートに従って走行する前に、いくつかの遅いターンを行う初期化プロセスを持つ追加バージョンを各データレコードに提供する。

2) 知覚タスク: ループCの運転記録から、深度画像、オリジナルフレーム、およびそれらに対応するセマンティックセグメンテーションのグランドトゥールズを生成する。キャプチャされたフレームは、トレーニングセット、検証セット、テストセットとして、7:1:2の割合で3つのサブセットに分割される。訓練セットは知覚モデルの訓練にのみ使用され、検証セットは訓練セットの補足として、またはモデル選択に使用することができ、テストセットはモデル



(a) Front view



(b) Bird-eye view

図7. 正面図とBEVにおける画像とセマンティックセグメンテーションのグランドトゥールズの比較。シーン内のアイテムは別々にラベル付けされている: 走行可能エリア、壁、柱、静的車両、駐車線、車線、衝突回避ストリップ、スピードバンプ、矢印。左列は正面図の元画像と投影されたBEV、右列はグランドトゥールズを示す。

評価のみである。

D. データセットの内容

データセットをrosvbag形式で記録する。ステレオカメラとその周辺カメラは、他のセンサーと部分的に作動する。なぜなら、長い期間と量のセンサーは、メモリと計算にコストがかかるからである。しかし、ループA、Bで車両が走行する際に、サブセクションIII-Bで説明したすべてのセンサーが作動するバッグを用意する。レコードはRoute SpeedLimit InitializeProcess Sensor.bagと命名する。

- ルートには、バッグファイルがどのループに記録されているかが記述されている。候補はループA、ループB、ループC。
- SpeedLimit は平均移動速度を表す。候補者は速く、遅い。
- InitializeProcess は、IMU の初期化に役立つ余分な処理があるかどうかを記述する。候補者が妨害されている(確かに)、直接的である(ない)。
- センサーは、記録中に作動するセンサーを記述する。ステレオとは、ステレオカメラやその他のセンサーが作動すること。周囲とは、魚眼カメラやその他のセンサーが作動すること。フルとは、すべてのセンサーが作動することを意味する。

セマンティックセグメンテーション、駐車場検出、奥行き推定など、サポートされる知覚タスクのために、5,000以上のフレームがグランドトゥールズとして提供される。いくつかのサポートタスクの詳細はセクションIVで紹介する。図7は、本データセットにおける分類ラベルの例である。

IV. EXPERIMENTS

A. 意味的セグメンテーション

意味的セグメンテーションは標準的な知覚タスクであり、多くの手法やネットワークがこの問題を解決する。

TABLE III

Cモデル間のMiouとFPSの比較.

Model	MIOU	FPS
BiSeNet [24]	0.8531	62.59
SFNet [25]	0.9091	6.45

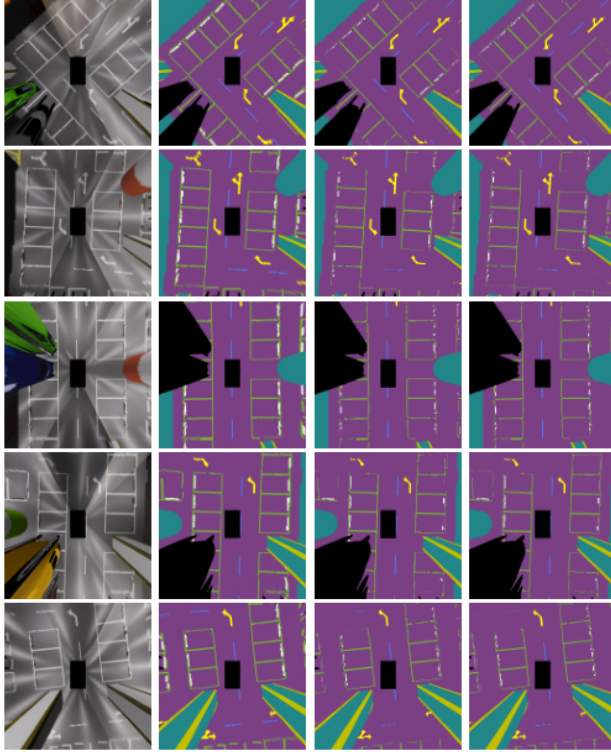
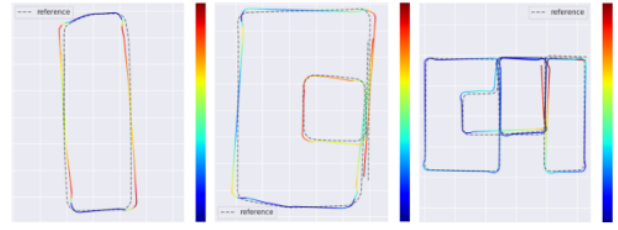


図8. 左側から右へ、1列目は元のBEV画像、2列目はセグメンテーションのグラントゥールズ、3列目はBiSeNet[24]の推論結果、最後の列はSFNet[25]の推論結果である。

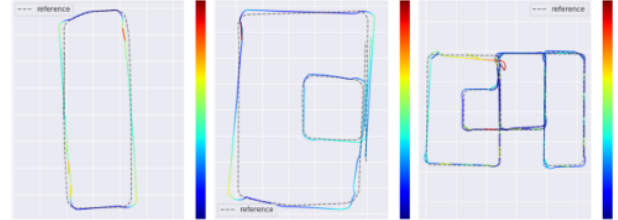
しかし、自律走行利用においては、このタスクの焦点は、精度と効率のトレードオフに焦点を当てている。BiSeNet[24]は、空間経路と文脈経路を含むバイラテラルセグメンテーションネットワークであり、精度を確保しながら効率化を実現している。SFNet[25]は、特徴量融合を実現するフローライメントモジュールを提案し、Cityscape[2]テストセットで高いmIoUを達成した。両ネットワークの性能を評価し、我々のデータセットでアルゴリズムがどの程度機能するか、また、同時使用の実用性をテストする。結果はTab. IIIと図8で見ることができる。

B. 視覚SLAM

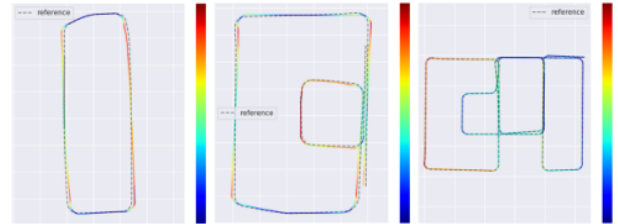
ORB-SLAMシリーズはオープンソースの最先端Visual-SLAMソリューションであり、ORB-SLAM3 [26]はこのシリーズの最新研究で、堅牢な性能を持つ。VINSFusion [27]も古典的な特徴ベースのVisual-SLAM手法である。我々のデータセットにおいて、3つのルートで頑健性と精度をベンチマークした。実験結果をTab. IVと図9に示す。EVO[28]を用いて絶対軌跡誤差(ATE)を評価する。



(a) ORB-SLAM3 [26]



(b) VINS-Fusion [27]



(c) LIO-SAM [29]

図9. Visual SLAMとLiDARに基づくSLAM手法の3つのルートでの評価。左列はループA、内側列はループB、右列はループCで評価される。

TABLE IV

C3つの経路におけるスラムアルゴリズムのアベ誤差の比較.

SLAM	Route	Max	Median	Min	RMSE	STD
ORB3 [26]	A	2.70	1.86	0.17	1.78	0.77
	B	3.12	1.88	0.23	1.96	0.78
	C	10.25	3.14	1.14	4.78	2.57
VINS-F [27]	A	4.11	1.78	0.19	1.70	0.71
	B	6.28	1.82	0.06	1.95	0.77
	C	7.30	2.71	1.75	3.03	0.99
LIO-SAM [29]	A	1.52	0.87	0.04	0.92	0.38
	B	1.40	0.71	0.03	0.80	0.32
	C	1.37	1.01	0.80	1.04	0.14

C. LiDARベースSLAM

Visual-SLAMアルゴリズムはより豊富な情報を持ち、必要リソースも少ないが、LiDARベースのSLAMは信頼性の高い精度とロバスト性を持つ古典的な手法である。LIO-SAM [29]は、IMUとGPS情報を持つ最近の密結合LiDAR SLAM手法である。LIO-SAM[29]の結果を評価し、その結果を図9と表IVに示す。IV.

V. CONCLUSIONS

我々は、完全なセンサーを持つ地下駐車シナリオにおける自律走行タスクをサポートする、新しい知覚とSLAMベンチマークであるSUPSデータセットを提供する、

例えば、サラウンドフィッシュアイカメラ、フォワードピンホールカメラ、デプスカメラ、LiDAR、GNSS、IMUなど。提案データセットを用いて、セマンティックセグメンテーション、駐車場検出、SLAMアルゴリズムを評価し、その実用性を検証する。SUPSデータセットに加え、仮想シーンと車両のセットアップをシミュレータ上でオープンソース化し、特定の適応を行うユーザーを対象とする。BEV視界の路側に沿って完全な駐車スペースをカバーできる大きなFovを提供したので、駐車スロットの特別な記述を提案した。今後の方向性の一つは、駐車スロットを記述するために低レベルのコーナーやセンターラインを使用する代わりに、オブジェクトレベルの検出方法を開発することである。物体レベルの物体検出は、その後の運動計画や意思決定手順において、散乱構造よりも有用である可能性がある。

ACKNOWLEDGMENT

本研究は、上海市科学技術専攻プロジェクト(No.2018SHZDZX01)、ZJラボ、上海脳科学・脳刺激技術センターの支援を受けている。

REFERENCES

- [1] J. Huang, L. Zhang, Y. Shen, H. Zhang, S. Zhao, and Y. Yang, "DMPPS: A novel approach for parking-slot detection using directional marking-point regression," in *2019 IEEE International Conference on Multimedia and Expo (ICME)*, Jul. 2019, pp. 212–217.
- [2] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [3] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stückler, and D. Cremers, "The tum vi benchmark for evaluating visual-inertial odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1680–1687.
- [4] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.
- [5] S. Yogamani, C. Hughes, J. Horgan, G. Sistu, P. Varley, D. O'Dea, M. Uricár, S. Milz, M. Simon, K. Amende, *et al.*, "Woodscape: A multi-task, multi-camera fisheye dataset for autonomous driving," *arXiv preprint arXiv:1905.01489*, 2019.
- [6] K. Minoda, F. Schilling, V. Wüest, D. Floreano, and T. Yairi, "Viode: A simulated dataset to address the challenges of visual-inertial odometry in dynamic environments," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1343–1350, 2021.
- [7] J. K. Suhr and H. G. Jung, "Automatic parking space detection and tracking for underground and indoor environments," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 9, pp. 1–1, 2016.
- [8] X. Li, "Gnss repeater based differential indoor positioning with multi-epoch measurements," *IEEE Transactions on Intelligent Vehicles*, pp. 1–1, 2021.
- [9] B. Shin, J. H. Lee, C. Yu, C. Kim, and T. Lee, "Underground parking lot navigation system using long-term evolution signal," *Sensors*, vol. 21, no. 5, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/5/1725>
- [10] C. Min, J. Xu, L. Xiao, D. Zhao, Y. Nie, and B. Dai, "Attentional graph neural network for parking-slot detection," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3445–3450, 2021.
- [11] J. Philion and S. Fidler, "Lift, splat, shoot: Encoding images from arbitrary camera rigs by implicitly unprojecting to 3d," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Cham: Springer International Publishing, 2020, pp. 194–210.
- [12] F. Bieder, S. Wirges, J. Janosovits, S. Richter, Z. Wang, and C. Stiller, "Exploiting multi-layer grid maps for surround-view semantic segmentation of sparse lidar data," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 1892–1898.
- [13] G. Rong, B. H. Shin, H. Tabatabaee, Q. Lu, S. Lemke, M. Možeiko, E. Boise, G. Uhm, M. Gerow, S. Mehta, *et al.*, "Lgsvl simulator: A high fidelity simulator for autonomous driving," *arXiv preprint arXiv:2005.03778*, 2020.
- [14] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The euroc micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 2016.
- [15] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nusenes: A multimodal dataset for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [16] X. Huang, X. Cheng, Q. Geng, B. Cao, D. Zhou, P. Wang, Y. Lin, and R. Yang, "The apolloscape dataset for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.
- [17] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "Bdd100k: A diverse driving dataset for heterogeneous multitask learning," 2018.
- [18] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2021.
- [19] J. Chen, S. Chen, M. Bai, J. Pu, J. Zhang, and J. Gao, "Graph decoupling attention markov networks for semisupervised graph node classification," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2022.
- [20] M. Kaneko, K. Iwami, T. Ogawa, T. Yamasaki, and K. Aizawa, "Mask-slam: Robust feature-based monocular slam by masking using semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.
- [21] B. Bescos, J. M. Fácil, J. Civera, and J. Neira, "Dynaslam: Tracking, mapping, and inpainting in dynamic scenes," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4076–4083, 2018.
- [22] "Unity," <http://www.unity3d.com>.
- [23] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [24] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, "Bisenet: Bilateral segmentation network for real-time semantic segmentation," 2018. [Online]. Available: <https://arxiv.org/abs/1808.00897>
- [25] X. Li, A. You, Z. Zhu, H. Zhao, M. Yang, K. Yang, and Y. Tong, "Semantic flow for fast and accurate scene parsing," in *ECCV*, 2020.
- [26] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel, and J. D. Tardós, "Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874–1890, 2021.
- [27] T. Qin, S. Cao, J. Pan, and S. Shen, "A general optimization-based framework for global pose estimation with multiple sensors," 2019.
- [28] M. Grupp, "evo: Python package for the evaluation of odometry and slam," <https://github.com/MichaelGrupp/evo>, 2017.
- [29] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5135–5142.