# LiDAR データの距離・反射強度・欠損分布を表現する

# 深層生成モデルの開発

○劉 瀟文, 中嶋 一斗, 宮脇 智也(九州大学),

岩下 友美(NASA ジェット推進研究所), 倉爪 亮(九州大学)

## **Development of Deep Generative Models**

## for LiDAR Range, Reflectance, and Raydrop Distributions

O Xiaowen LIU, Kazuto NAKASHIMA, Tomoya MIYAWAKI (Kyushu University),

Yumi IWASHITA (NASA Jet Propulsion Laboratory), and Ryo KURAZUME (Kyushu University)

Abstract: In recent years, LiDAR simulators have increasingly attracted attention for training machine learning models for scene understanding. However, discrepancies between simulated and real-world environments can impair model performance. To mitigate the issue, this paper introduces a deep generative model for LiDAR domain adaptation built upon generative adversarial networks. Our approach jointly learns the distributions of LiDAR range, reflectance, and raydrop. We demonstrate its generation fidelity and diversity on the KITTI-360 dataset.

## 1. はじめに

LiDAR センサは、パルスレーザを用いて全周囲の距離を測定することで、周囲環境の 3D 点群を得ることができる。一般的に、深度カメラやレーダと比べて視野角が広く、環境照度によらずロバストに計測できるため、自動運転や自律移動ロボットに不可欠な構成要素の一つとなっている。LiDAR センサから得られる 3D 点群は、SLAM 等の自己位置推定だけでなく、機械学習に基づく障害物検知に広く用いられる。

-方で,機械学習に基づく障害物検知においては, モデルを構築するための大規模学習データのアノテー ションコストが課題となっている.特に LiDAR 点群 は、一つのスキャンに含まれる点数が多く、3D 空間の 整合性を保ちながらアノテーションを行う必要がある ため、他のドメインに比べてスケーラビリティが低い. この問題に対して、自動運転シミュレータ上で擬似 的な LiDAR スキャンと正解ラベルの対を自動合成し、 機械学習モデルの学習データとして利用する Sim2Real アプローチが注目されている.一方,物体の状態・材 質に由来するレーザの複雑な反射特性の再現が難しい ため、シミュレーションデータには実データ特有の欠 損ノイズや反射強度が含まれず、汎化性能が低下する 問題がある.そのため、実環境に適応する学習データ を自動合成するためには、シミュレーションデータに 対して反射特性モダリティを復元する Sim2Real 手法が 必要である.本稿では、LiDAR データの Sim2Real を目 的として、実環境の LiDAR データを構成する距離値・ 反射強度・欠損ノイズを深層生成モデルとしてモデリ ングする手法を紹介する.

## 2. 関連研究

近年,多層ニューラルネットワークによって生成モ デルを構築するフレームワークが数多く提案されてい る.本研究では,最も成功したフレームワークの一つ である敵対的生成ネットワーク(GAN)<sup>[1]</sup>を採用す る.GAN は,画像・音声・テキストなどの様々なド メインのデータ生成において実証されている. LiDAR データへの適用については, Caccia ら<sup>[2]</sup> が自然画像分 野で実証された DCGAN<sup>[3]</sup> を用いて LiDAR 距離画像 の生成モデルを構築している.また, Nakashima ら<sup>[4]</sup> は, LiDAR 距離画像に含まれる離散的な欠損ノイズが 生成品質低下に影響することを指摘し, 欠損ノイズを 自動的に分離表現する DUSty V1 を提案している.さ らに, Nakashima ら<sup>[5]</sup> は, LiDAR 距離画像の GAN に ニューラル場表現を導入することで,任意の角度解像 度でデータ表現できる DUSty V2 を提案している.上述の先行研究では,距離値および欠損分布の生成に成 功しているものの,シーン理解に重要な反射強度の学 習には至っていない.本研究では, DUSty V1・V2を拡 張し,距離・反射強度・欠損分布の学習を目的とする.

#### 3. 提案手法

#### 3.1 敵対的生成ネットワーク(GAN)

GAN は、生成器 (Generator) と識別器 (Discriminator) と呼ばれる二つの多層ニューラルネットワークによっ て構成される.生成器は潜在変数からデータを生成し、 識別器は入力されるデータを実データ・生成データの いずれかに識別する.学習時は、生成器は識別器を欺 くように、識別器は実データ・生成データを正しく識 別するように、それぞれのモデル重みを更新する.両 者が互いに競合して学習することで、JS ダイバージェ ンス最小化のもと生成器が出力する生成分布はデータ 分布に近づいていく.

LiDAR データを対象とした GAN である DUSty v1<sup>[4]</sup> および DUSty v2<sup>[5]</sup> について述べる.距離画像に含まれ る欠損ノイズは離散的な二値ノイズであるため,連続 関数によって構成される GAN の学習を不安定にする 可能性がある. DUSty v1<sup>[4]</sup> では,欠損を含む距離画像 を直接生成する代わりに,全画素の距離値と対応する 欠損確率による 2 種類の中間表現を生成する.次に, 欠損確率を生起確率としたベルヌーイサンプリングに よって二値ノイズを生成することで,欠損ノイズを重 畳した生成画像を得る.以上の生成画像をもとに敵対



Fig.1 Examples of KITTI<sup>[7]</sup> training samples



Fig.2 Logarithmic representation of images

的学習を行うことで,生成器は連続的な距離値・欠損 確率による分解表現をもとに安定に学習することがで きる.

DUSty v2<sup>[5]</sup>は、座標入力によって対応値を回帰する ニューラル場表現<sup>[6]</sup>を導入することで、DUSty v1 で 提案された距離値・欠損確率の分解表現における空間 解像度を可変にしたものである.DUSty v2 の生成器で は、潜在変数に加えて、各画素に対応するレーザの照 射角度 (仰角、方位角)を入力し、対応する距離値およ び欠損確率を出力する.代表的な KITTI データセット を用いた実験では、距離画像の品質・多様性において DUSty v1 を上回る性能を示している.

#### 3.2 反射強度モダリティの追加

本研究ではまず, DUSty v1/v2 の距離値・欠損確率 の出力に反射強度を追加する.反射強度は,物体の材 質や形状に由来する特徴を持っており,物体検出やセ グメンテーション等の意味理解タスクに重要である. 図1に, KITTI データセット<sup>[7]</sup>における距離画像と反 射強度画像の例を示す.

#### 3.3 距離画像の対数表現

距離画像のダイナミックレンジは少数の遠距離画素 が支配的になり、シーン全域の精細な学習が困難とな る可能性がある.本研究では、先行研究<sup>[8]</sup>の対数距離 表現を GAN に適用することで、画像を多く占める近距 離画素のダイナミックレンジ改善を図る.図2に、元 の距離表現・逆深度表現・対数距離表現を比較する.

#### **3.4** 欠損分離に関する制約

前章で述べた対数距離表現を DUSty v2 に適用した場合,距離出力と欠損出力の分離学習が不安定となる挙動が見られた.この問題を解決するために,距離値出力xのうち欠損値 (x = 0)付近の画素について次のヒンジ損失  $\mathcal{L}_{mask}$ を導入することで分離を補助する.

$$\mathcal{L}_{\text{mask}} = \max\left(0, -x + \epsilon\right),\tag{1}$$

ただし、 $\epsilon$  は許容誤差を決めるハイパーパラメータである.  $\mathcal{L}_{mask}$  は、生成器学習時の敵対的損失とともに最小化する. 図 3 に  $\mathcal{L}_{mask}$  の有無による距離値出力の例を示す.



Fig.3 Effect of the mask regularization term  $\mathcal{L}_{\rm mask}$ 

手法	反射強度	対数距離	$\mathrm{FPD}\downarrow$	$MMD\downarrow$
DUSty V1 DUSty V1 DUSty V1 DUSty V1	$\checkmark$	√ √	$241.3 \\ 84.8 \\ 818.1 \\ 107.0$	$42.6 \\ 5.3 \\ 136.2 \\ 13.2$

Table 1 CNN based Gan

## 4. 実験

### 4.1 データセット

本研究では,評価実験に KITTI データセット<sup>[7]</sup> を用 いる. KITTI データセットは,ドイツの都市部や高速 道路などの複数のシーンにおいて,カメラ・LiDAR・ GPS データを記録したものである.本研究では,座標 集合で与えられている LiDAR データを 64 段階の仰角 に応じて分割した後,それぞれを 512 点に間引くこと で 64×512 の距離画像・反射強度画像を作成した.

#### 4.2 評価指標

生成データは、品質および多様性の観点から FPD<sup>[9]</sup> と MMD<sup>[5]</sup> によって評価する.FPD は、自然画像の生 成モデル評価において標準的に使用される FID を、 点群に適応させた指標である.本研究では、まず実 データ・生成データの距離画像を学習済み PointNet<sup>[9]</sup> に入力し、中間特徴を抽出する.次に、特徴空間上の 実データ・生成データの集合間で Fréchet 距離を算出 する.MMD は、同様の特徴空間上において Squared maximum mean discrepancy を計算する.

#### 4.3 定量的評価

表1にDUSty v1をベースとした提案手法の結果を示 す.距離画像に対数表現を導入すると、元のベースラ インに対して各スコアが大幅に改善することを確認で きる.また、反射強度を追加したモデルはスコアが大 きく低下することがわかった.図1で示されるように、 距離画像と反射強度画像はテクスチャおよび輝度値の 分布が大きく異なっていることが要因の一つとして考 えられる.対数距離とともに反射強度を学習すること で、対数距離のみの場合と同等の結果を得ることがで きた.

表2にDUSty v2をベースとした提案手法の結果を示 す.表1のDUSty v1と異なり,対数距離の導入によっ てスコアが大きく低下した.第3.4章で述べた欠損分 離に関する制約を導入すると,元のベースラインから FPD が向上し,MMD は同等に抑えることができた. 最後に,反射強度を導入した結果では,距離画像のみ の結果と比べて劣るものの同等の結果を得ることがで

手法	反射強度	対数距離	欠損制約項	$\mathrm{FPD}\downarrow$	$MMD\downarrow$
DUSty V2 DUSty V2 DUSty V2 DUSty V2	~		$\checkmark$	93.8 294.0 78.5 98.1	3.8 38.1 7.8 9.4

Table 2 Neural field based Gan



Fig.4 Generated examples of DUSty V1-based methods



Fig.5 Generated examples of DUSty V2-based methods

## きた.

## 4.4 定性的評価

図4に対数距離・反射強度学習を導入した DUSty v1 による生成データの例,図5に対数距離・欠損分離制 約・反射強度学習を導入した DUSty v2 による生成デー タの例を示す.いずれの手法においても,距離画像・ 欠損ノイズ・反射強度のそれぞれについて,実データ の特徴を詳細に表現できていることがわかる.

## 5. まとめと今後の予定

本稿では、LiDAR データの Sim2Real 応用を目的とし て、実データの距離値・欠損ノイズ・反射強度の分布 を表現する深層生成モデルを紹介した.特に、敵対的 生成ネットワークに基づく先行研究 DUSty v1/v2 に対 して、対数距離表現・分離表現の制約項を導入するこ とで、ベースラインと同程度のスコアを維持しながら 反射強度画像の生成を可能にすることができた.今後 は、生成した反射強度画像の品質・多様性評価、学習 した GAN を利用した Sim2Real タスク評価に取り組む.

## 謝辞

本研究の一部は, JSPS 科研費 JP23K16974, JST 【ムーンショット型研究開発事業】 グラント番号 【JPMJMS2032】の支援を受けたものである.

## 参考文献

- Ian Goodfellow et al.: "Generative adversarial nets". Proceedings of the Advances in Neural Information Processing Systems (NeurIPS). 2014, pp. 2672–2680.
- [2] Lucas Caccia et al.: "Deep Generative Modeling of Li-DAR Data". Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2019, pp. 5034–5040.
- [3] Alec Radford, Luke Metz, and Soumith Chintala: "Unsupervised representation learning with deep convolutional generative adversarial networks". *arXiv preprint arXiv:1511.06434*, (2015).

- [4] Kazuto Nakashima and Ryo Kurazume: "Learning to Drop Points for LiDAR Scan Synthesis". *Proceedings* of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2021, pp. 222–229.
- [5] Kazuto Nakashima, Yumi Iwashita, and Ryo Kurazume: "Generative Range Imaging for Learning Scene Priors of 3D LiDAR Data". *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 2023, pp. 1256–1266.
- [6] Ivan Skorokhodov, Savva Ignatyev, and Mohamed Elhoseiny: "Adversarial Generation of Continuous Images". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021, pp. 10753– 10764.
- [7] Andreas Geiger et al.: "Vision meets robotics: The KITTI dataset". *The International Journal of Robotics Research* (*IJRR*) 32.11, pp. 1231–1237, (2013).
- [8] Vlas Zyrianov, Xiyue Zhu, and Shenlong Wang: "Learning to Generate Realistic LiDAR Point Clouds". Proceedings of the European Conference on Computer Vision (ECCV). 2022, pp. 17–35.
- [9] Dong Wook Shu, Sung Woo Park, and Junseok Kwon: "3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions". *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019.