LiDAR 点群アップサンプリングのための 欠損と反射率を考慮した深度マップ超解像

Qiongyu Li^{1,a)} 中嶋 一斗^{2,b)} 倉爪 亮^{2,c)}

概要

本稿では、2D 畳込みニューラルネットを用いた 3D Li-DAR 点群のアップサンプリング手法を提案する.提案手 法は、レーザ測距で得られる正距円筒図形式の深度・反射 率マップに対し、点欠損を考慮した 2D 畳込みによる超解 像処理を適用することで、高密度点群を得る.また、2D 表現と 3D 表現の両面で超解像プロセスを学習する損失関 数を提案する.実験では、屋外環境を対象とした 3D 点群 データセットを用いて提案手法の有効性を検証した.

1. はじめに

LiDAR センサは、自動運転における周囲環境計測, 歴 史的建造物のデジタルアーカイブなど、3D 構造を高精度 に計測する上で広く利用されている.最も一般的な計測方 式は、全方位レーザ照射によって各仰俯角・方位角の距離 値を逐次計測するもので、対象物の照明環境によらず安定 して広域点群を取得することができる.一方で、点密度に よって計測に要する時間は異なるだけでなく、一般的に高 密度 LiDAR センサの導入コストは高い.このような問題 の解決策として、LiDAR 点群のアップサンプリング技術 の開発は重要な課題である.

2. 関連研究

2.1 LiDAR データの表現方法

近年,点群等の 3D データを対象とした深層学習技術に ついて活発に議論されている.特に,点群データの主な表 現方法としては点集合,ボクセル,2Dマップの三つに分類 され,データ処理のアプローチは異なる.点集合表現は, 計測対象の 3D 構造を最も自然に表現しており,各点の 3D 座標値を順不同に並べた形式である.点集合を直接処理す る汎用アーキテクチャ [9] も提案されているが,処理単位 が小さく,LiDAR データのような大規模点群の学習には大

- ² 九州大学
- ^{a)} liqiongyu@irvs.ait.kyushu-u.ac.jp



図1 提案手法の処理フロー

きな計算コストとメモリを要する.ボクセル表現は,点群 を固定グリッドに分割し,グリッドごとに符号化する方式 である.構造化されていない点集合表現と異なり,近傍点 間の特徴を抽出することができるが,3D 演算に伴うメモリ 使用量が非常に高い.一方,一度の計測で得られる LiDAR 点群は,一対一対応の 2D マップとして簡易表現でき,計 算コストの少ない 2D 畳込み演算で処理できる利点がある. 例えば,Liら [6] は,円筒座標系で表現した屋外の LiDAR 点群を用いて,高精度な車体検出を実現している.

2.2 点群のアップサンプリング

点群のアップサンプリングとは、ノイズや不均一性を持 つ疎な点群の密度を上げる処理である.点群全般を取り 扱った例としては、Liら [7]の PU-GAN が挙げられる. PU-GAN では、点群データを直接処理する PointNet [9]を ベースに、均一で密な点集合を生成する敵対的生成ネット ワークを構成している.LiDAR 点群に対する結果を報告 しており、車や人の局所領域に対して高精度なアップサン プリングを実現しているが、一度に処理できる点集合が小 さく、路面勾配等の広域な構造に対処できていない.

2.3 画像超解像

本研究では,LiDAR 点群の2Dマップ表現に着目し,画 像超解像技術の応用を試みる.一般自然画像に対する超解 像処理では,CNNを用いて低解像度画像の入力から高解

¹ 大連理工大学

b) k_nakashima@irvs.ait.kyushu-u.ac.jp
c) kurazume@ait.kyushu-u.ac.jp

像度画像を推定するアプローチが主流である.代表的な手 法として,Ledig ら [5]の SRGAN が挙げられる.SRGAN では,複数の残差学習ユニットと Sub-Pixel Convolution による特徴マップ変換で高解像度画像を出力する.また, 出力する画素ごとの一致度と出力画像全体の大局的整合度 を基にパラメータ学習を行う.

3. 提案手法

本手法では, Ledig ら [5] の SRGAN をベースとした深度 マップの超解像生成器と敵対的生成学習のための識別器を 構成する.さらに, 畳込み演算の点欠損対応, マルチモー ダル融合層の追加, 損失関数の改良を行う.超解像の拡大 率は SRGAN と同様に4とする.その他詳細なハイパーパ ラメータ, モデル構造は論文 [5] を参照されたい.

3.1 入力データと前処理

垂直 H ラインおよび水平 W 点のスキャンを行う LiDAR センサでは H×W 組の距離値と反射率が得られる.同時 に,これらの計測点を縦軸を仰俯角,横軸を方位角とし た 2D グリッドに投影すると、距離値・反射率を各画素の チャンネルとした (H,W,2) の深度・反射率マップで表現 できる.反射率はレーザ照射対象ごとに連続性を有してお り、深度マップの形状特徴抽出を補助できる可能性が高 い. 一方, レーザ反射を集光できない空や鏡面では計測点 の欠損が生じる.特に、街路樹のように微細な欠損が集ま る領域では、高解像度データの深度チャンネルに多くの不 連続なテクスチャが生じ、モデル学習を困難にすること が確認された. そこで, まず入力データに対して, Jason ら [4] の深度マップ欠損補間手法 IP-Basic*1を適用する. さらに,計測点の有無を示す二値の欠損マスクを新たに追 加し, (H,W,3)の入力データを構成する. 深度チャンネル の欠損領域は 0m を設定し,深度と反射率の値域をそれぞ れ [-1,1] に正規化する.

3.2 点欠損を考慮した畳込み演算

深度マップ中の欠損領域の影響を抑制するために,全て の畳込み層をLiuら [8]のPartial Convolution に置き換え る.Partial Convolution は,画像欠損修復タスクの解法と して提案されたものである.畳込みカーネル内の有効画素 のみから応答値を算出し,その画素数に応じて正規化する ことで,欠損の広がりや有無に対して特徴抽出の一貫性を 与えることができる.入力データの有効領域を表すマスク として,3.1節の欠損マスクを利用する.このマスクは層 伝播に従って更新され,SRGAN 後段のアップサンプル層 では最近傍補間によるマスク拡大を行う.

3.3 深度情報と反射率情報の融合

第1畳込み層を深度マップと反射率マップの二つの経路 に分離し,出力された特徴マップをチャンネル結合する. チャンネル結合後は,SRGANと同様の単一伝播路により 超解像処理を行う.特に,第1層では深度マップに加えて, 点欠損マスクが入力される.

3.4 損失関数

SRGAN と同様に、画素ごとの一致を測る L1 損失 *L*_{L1} と、画像全体の整合性を測る敵対的損失 *L*_{GAN} を採用する. L1 損失は、生成された超解像画像と高解像度画像の対応画 素間で絶対誤差を算出する.本手法においては、LiDAR セ ンサがレーザ照射する計測点ごとの距離誤差を意味する. 一方、敵対的損失は、実画像と生成画像を識別する識別器 の出力値によって与えられ、生成器との min-max ゲーム で最小化される.本手法では、敵対的損失の定式化として、 SRGAN 同様の Non-Saturating GAN [2] を採用した.

一方, $\mathcal{L}_{L1} \geq \mathcal{L}_{GAN}$ のみを用いた予備実験では, 2Dマッ プ上の形状が近似できているものの,再構成された点群上 では物体のない領域で多くのノイズ点が現れることが分 かった.そこで,上記二つの損失に加え, 3D 座標上の一致 度を評価する Chamfer 損失 [1] を導入する.超解像で得ら れる生成画像の各画素は,極座標系における特定の仰俯角・ 方位角に対応しており,画素値のみを変数として 3D 座標 に変換することができる.この再構成された 3D 点群間で Chamfer 距離(CD)を算出し,損失 \mathcal{L}_{CD} として最小化す る.Chamfer 距離は,二つの点群間で双方向に最近傍点を 探索し,対応点間の L2 距離を算出する.最終的に,損失 ごとの係数 λ_i を用いて以下の重み付き損失を最小化する.

$$\mathcal{L} = \lambda_1 \mathcal{L}_{L1} + \lambda_2 \mathcal{L}_{GAN} + \lambda_3 \mathcal{L}_{CD} \tag{1}$$

4. 実験

4.1 データセット

高密度 LiDAR センサで計測された点群データを用い て,提案手法の有効性を検証する.データセットは,Jung ら [3] の Multi-modal Panoramic 3D Outdoor (MPO) デー タセット*²を用いた.MPO データセットは,車体上部に 取り付けた高密度 LiDAR センサ FARO Focus 3D で 650 箇所の屋外環境を計測したもので,各スキャンマップは垂 直解像度 1,757 と水平解像度 5,140 を持つ.また,各計測 点は 120m までの距離値と [0,1] の反射率を持つ.本研究 では,これらの高解像度マップから低解像度マップを合成 し,提案手法によって超解像化を行う.実験に用いた計算 機の制限から,まず正解となる全ての高解像度マップを最 近傍法により 1/4 に縮小した.計 650 個のスキャンのう ち,554 枚を学習データ,96 枚をテストデータとした.

*2 http://robotics.ait.kyushu-u.ac.jp/kyushu_datasets/

^{*1} https://github.com/kujason/ip_basic

4.2 学習と評価方法

超解像モデルの学習では、まず高解像度マップから 384×384のパッチ画像をランダムクロップする.さらに、 最近傍法による 1/4 解像度の 96×96 画像を合成し、パッ チベースの超解像学習を行った.式1の係数 $\{\lambda_1, \lambda_2, \lambda_3\}$ は、それぞれ $\{1, 10^{-3}, 10^2\}$ を設定した.また、ミニバッ チサイズを 8、反復回数を 200,000、学習係数を 10^{-4} に固 定し、Adam によるパラメータ更新を行った.評価する場 合は、高解像度マップ上を固定ストライドで走査しながら、 各パッチの結果を統合する.実験では、ベースラインであ る SRGAN ($\mathcal{L}_{L1} \geq \mathcal{L}_{GAN}$ のみ)と三つの数値補間法(最 近傍、バイリニア、バイキュービック)を比較手法とし、ス キャンマップ上の距離誤差として平均二乗誤差(MSE)と ピーク信号対雑音比(PSNR)、画像としての生成品質を測 る SSIM、再構成された点群における誤差として Chamfer 距離(CD)を算出した.

4.3 結果

結果を表1に示す.表1では,提案手法の要素ごとの効果 を調査するために,手法の一部を部分的に除去した結果とし て提案手法A・B・Cを比較している.MSE, PSNR, SSIM に関しては提案手法が最も良いスコアを示したが, Chamfer 距離に関しては最近傍法に劣る.最近傍法の Chamfer 距 離が非常に小さい原因として,数値補間される点が近傍点 から選択されるため,局所的な分散が小さいことが考えら れる.対応点探索に最近傍法を用いる Chamfer 距離の性 質上,実験で用いた広域点群の遠距離領域でこの分散の影 響は大きい.提案手法A・B・C の Chamfer 距離に着目す ると,最良ではないものの,Chamfer 損失による点群の直 接最適化の効果が確認できる.

図2に,最近傍法,バイリニア法,提案手法Cで生成された点群の例を示す.点群の3D構造を見やすくするために,生成点群上で30近傍点による法線推定[10]を行い,その向きに従って配色している.Chamfer距離において最良であった最近傍法の例を見ると,局所的に分散が小さいものの勾配構造を捉えておらず,道路のような平面がステップ状になっている.バイリニア法では,物体のない領域にノイズ点が頻出している.提案手法では,ノイズ点が比較的少なく,滑らかなアップサンプリングが実現できている.

5. まとめと今後の課題

本稿では、LiDAR センサのレーザ測距で得られる深度・ 反射率マップに対して超解像を適用することで、点群アッ プサンプリングを行う手法を提案した.また、深度・反射 率マップを入力として点欠損を考慮しながら超解像を行う 畳込みニューラルネット、および 2D マップ上の一致度と 3D 点群上の一致度を組み合わせた損失関数を提案した. 屋外環境の 3D 点群データセットを用いた実験では、生成 された高解像度画像上の評価および再構成点群上の評価で その効果を確認した.今後の課題として,正距円筒図表現 に伴う歪みの対応.センサ近傍領域におけるノイズ抑制が 挙げられる.

謝辞

本論文は JSPS 科研費 JP20H00230 の助成を受けた.

参考文献

- Fan, H., Su, H. and Guibas, L. J.: A point set generation network for 3d object reconstruction from a single image, *Proceedings of the IEEE Conference on Computer* Vision and Pattern Recognition (CVPR), pp. 605–613 (2017).
- [2] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative adversarial nets, *Advances in Neural Information Processing Systems (NIPS)*, pp. 2672–2680 (2014).
- [3] Jung, H., Oto, Y., Mozos, O. M., Iwashita, Y. and Kurazume, R.: Multi-modal panoramic 3D outdoor datasets for place categorization, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, pp. 4545–4550 (2016).
- [4] Ku, J., Harakeh, A. and Waslander, S. L.: In defense of classical image processing: Fast depth completion on the cpu, *Proceedings of the Conference on Computer* and Robot Vision (CRV), IEEE, pp. 16–22 (2018).
- [5] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. et al.: Photo-realistic single image superresolution using a generative adversarial network, *Proceedings of the IEEE Conference on Computer Vi*sion and Pattern Recognition (CVPR), pp. 4681–4690 (2017).
- [6] Li, B., Zhang, T. and Xia, T.: Vehicle detection from 3d lidar using fully convolutional network, *Robotics: Sci*ence and Systems (2016).
- [7] Li, R., Li, X., Fu, C.-W., Cohen-Or, D. and Heng, P.-A.: PU-GAN: A Point Cloud Upsampling Adversarial Network, *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, pp. 7202–7211.
- [8] Liu, G., Reda, F. A., Shih, K. J., Wang, T.-C., Tao, A. and Catanzaro, B.: Image inpainting for irregular holes using partial convolutions, *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 85–100 (2018).
- [9] Qi, C. R., Su, H., Mo, K. and Guibas, L. J.: Pointnet: Deep learning on point sets for 3d classification and segmentation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 652–660 (2017).
- [10] Zhou, Q.-Y., Park, J. and Koltun, V.: Open3D: A Modern Library for 3D Data Processing, arXiv:1801.09847 (2018).

示す.↓は低いほど望ましく,↑は高いほど望ましいことを示す.							
手法	入力	Partial Conv. [8]	Chamfer 損失 [1]	$\Big \text{ MSE } (\times 10^{-3}) \downarrow$	$\mathrm{SSIM}\uparrow$	$\mathrm{PSNR}\uparrow$	CD (×10 ⁻⁵) \downarrow
最近傍法	深度			6.04	0.896	23.6	9.12
バイリニア法	深度			4.77	0.904	24.7	19.9
バイキュービック法	深度			5.28	0.901	24.2	20.2
SRGAN [5]	深度			5.25	0.895	23.7	17.2
提案手法 A	深度	\checkmark		<u>4.69</u>	0.907	24.3	15.5
提案手法 B	深度	\checkmark	\checkmark	4.94	0.908	24.6	12.3
提案手法 C	深度・反射率	\checkmark	\checkmark	<u>4.00</u>	0.922	25.3	9.17

表 1 MPO データセット [3] を用いた超解像点群の精度評価.上位二つを下線,一位を太字で 示す.↓は低いほど望ましく,↑は高いほど望ましいことを示す.



図 2 MPO データセット [3] のテストセットを用いた生成結果