時系列平滑化を用いた列車前方映像に対するセマンティックセグメンテーション手法の検討

A Study on Semantic Segmentation of Railway Images using Temporal Smoothing

振津勇紀 [1] Yuki Furitsu
新村文郷 [1] Fumito Shinmura
出口大輔 [1] Daisuke Deguchi
川西康友 [1] Yasutomo Kawanishi
井手一郎 [1] Ichiro Ide
村瀬洋 [1] Hiroshi Murase
向嶋宏記 [2] Hiroki Mukojima
長峯望 [2] Nozomi Nagamine

名古屋大学 [1] Nagoya University
（公財）鉄道総合技術研究所 [2] Railway Technical Research Institute

1 はじめに

鉄道の営業距離は非常に長く，鉄道沿線設備の維持管理に多くのコストを要している．これに対し，LiDAR 等のレーザスキャナを用いて生成した 3D マップを用いることで，設備管理の省力化に向けた取り組みがなされている．しかし，LiDAR から得られる点群のみから設備の位置や種類を特定することは難しい．一方，運転中の列車外の状況把握に対する関心の高まりを受け，列車前方映像の蓄積が進められつつある．このような背景から，列車前方映像に画素単位でラベルを付与するセマンティックセグメンテーションを行なうことでラベル付きの映像を得たあと，SLAM 技術によって線路空間のラベル付き 3D マップを自動生成する技術開発を進めている．本発表では，3D マップ生成の前段階の処理として，移動物体が少ない鉄道環境に特化した時系列平滑化を用いた列車前方映像のセマンティックセグメンテーション手法を提案する．

2 提案手法

まず，列車前方映像を入力として，DeepLabv3+[2] により映像中の各フレームに画素単位でラベル付与を行なう．この際，セグメンテーション結果として画素毎にクラス尤度を得る．次に，フレーム間の Optical Flow を PWC-Net[3] を用いて計算し，計算したフローを基に各フレームとその前後フレームの画素を対応付ける．そして，セマンティックセグメンテーション結果と画素間対応付けを基に画素単位でラベルの時間方向の平滑化を行なう．この際，対応する画素のクラス尤度の平均を計算し，最も尤度が高いクラスを新たなラベルとする．

3 実験

提案手法の有効性を確認するため，列車前方映像を用いたセマンティックセグメンテーションとその平滑化に関する評価実験を行なった．

提案手法の学習と評価のため，手作業で画素単位のアノテーションを付与した列車前方映像のフレーム 9 枚を用意した．そして，そこから DeepLabv3+[2] の学習用に 8 枚，評価用に 1 枚のラベル付き画像を使用した．データ数の不足を補うため，Cityscapes[1] データセットで学習済みのモデルに対して転移学習を行なった．

次に，注目フレームとその前後複数フレームでセグメンテーション結果に対して平滑化処理を行なった．対象フレームとその前後を合わせた 3 フレーム $(N-1, N, N+1)$ を用いた際の平滑化結果を図 1 に示す．また，平滑化前後での 6 クラスの Intersection over Union（Class IoU）及び全クラスの平均（Mean IoU）の比較結果を表 1 に示す．この結果から，平滑化処理前と比べて処理後は Class IoU や Mean IoU が改善することを確認した．

表 1 平滑化処理前後での Class IoU 及び MIoU の変化

	建物	柵	柱	車両	軌条	軌道	MIoU
処理前	0.476	0.469	0.840	0.747	0.890	0.840	0.558
処理後	0.534	0.476	0.847	0.753	0.902	0.834	0.568

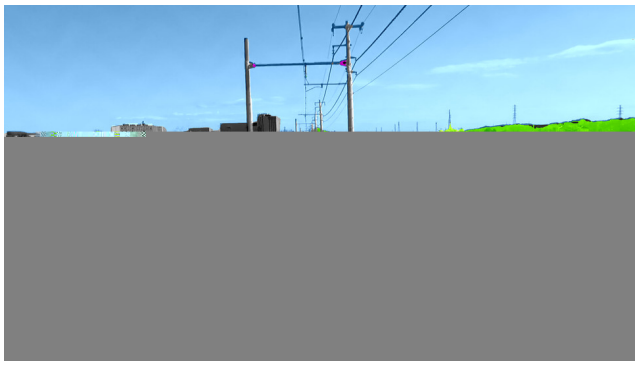

図 1 セグメンテーション結果（ピンク：軌条，赤：軌道）

4 むすび

時系列平滑化を用いた列車前方映像に対するセマンティックセグメンテーション手法を提案した．今後の課題として，平滑化結果の定量的な評価や，セグメンテーション結果からのラベル付き 3D マップ構築が挙げられる．

謝辞 本研究の一部は科学研究費補助金による．

参考文献

[1] M. Cordts et al., “The Cityscapes Dataset for Semantic Urban Scene Understanding,” CVPR, 2016

[2] L. Chen et al., “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” ECCV, 2018

[3] D. Sun et al., “PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume,” CVPR, 2018

Copyright © 2019 IEICE