

修士論文

画像認識による交差点付近における リアルタイム車両状態推定

Real-time vehicle status estimation
near intersections using image recognition

2022 年 12 月 20 日 提出

指導教官 岩井将行 教授

未来科学部 情報メディア学科

20FMI02 安齋凌介

修士論文要旨 2021 年度（令和3年度）

概要

交通工学はエッジセンサから得られたデータを活用することで、可能性が大きく広がる分野である。速度違反検知、ナンバープレート検知、渋滞検知などは日本のみならず先進国で導入されている。しかしながら、それらのデータを交通の最適化に活用している例はまだ少ない。これは、データの質の低さ、ラベルの欠如、データの利権などが障壁になっていることが考えられる。またそれらのセンサは大規模かつ高額になることが多く、交通量の多い交差点や高速道路に限定して設置されていることが多い。したがって、我々が普段使う道路や交差点ではこれらの導入は遅れていることが現状である。そこで、我々はオープンソースデータと安価なエッジセンサを活用し、交差点付近の車両の状態を推定することができるデバイスと、信号機を最適化するシステムを開発した。これにより今まで取ることができなかったその地点の詳細な車両のデータの取得と動的な信号機の信号操作を可能にした。

キーワード:

yolo, sort, リアルタイム, 状態推定

東京電機大学院未来科学研究科情報メディア学専攻

安齋 凌介

目次

第1章	序論	1
1.1	背景	2
1.2	本研究の目的	2
1.3	本論文の構成	2
第2章	関連研究	3
2.1	関連研究	4
第3章	提案手法	6
3.1	提案システム	7
3.1.1	概要	7
第4章	アルゴリズム	8
4.1	アルゴリズム概要	9
4.1.1	SORT について	9
4.1.2	射影変換	10
4.1.3	信号制御	10
第5章	結果	11
5.1	実験結果と考察	12
5.1.1	実験概要	12
第6章		13
	Acknowledgment	14

Off-campus Conference presentation	15
References	16
Appendix	18

图 目 次

表 目 次

第1章

序論

本章では，本研究の背景と目的および，本論文の内容構成について述べる．

1.1 背景

交通工学はエッジセンサから得られたデータを活用することで、可能性が大きく広がる分野である。速度違反検知、ナンバープレート検知、渋滞検知などは日本のみならず先進国で導入されている。しかしながら、それらのデータを交通の最適化に活用している例はまだ少ない。これは、データの質の低さ、ラベルの欠如、データの利権などが障壁になっていることが考えられる。またそれらのセンサは大規模かつ高額になることが多く、交通量の多い交差点や高速道路に限定して設置されていることが多い。したがって、我々が普段使う道路や交差点ではこれらの導入は遅れていることが現状である。そこで、我々はオープンソースデータと安価なエッジセンサを活用し、交差点付近の車両の状態を推定することができるデバイスと、信号機を最適化するシステムを開発した。これにより今まで取ることができなかったその地点の詳細な車両のデータの取得と動的な信号機の信号操作を可能にした。

1.2 本研究の目的

本研究は信号機の動的な制御を行うために、センサとして画像処理を用いて車両の詳細な挙動を認識することを目的としている。認識する項目は以下の通りである。車両の位置情報、車両の速度、(車両の種類、車両の軌跡、交差点内の車両の右左折待ち)これらのデータを用いることで、円滑な交通に繋がると考えている。

1.3 本論文の構成

本稿では、2章で関連研究について述べ、3章で提案システムの概要について説明し、4章ではアルゴリズムについて説明をする。5章では結果、6章では考察を述べる。最後に7章で今後の展望およびまとめを述べる。

第2章

関連研究

本章では、

2.1 関連研究

IoT (Internet of Things) の普及に伴い , さまざまな場所でのセンシングやデータの収集が可能となった . 交通工学の分野でも例外ではなく , 古くは車両データを扱った技術として速度違反自動取締装置 (オービス) が挙げられる . また , 近年では ITS (Intelligent Transport Systems : 高度道路交通システム) の観点からさらに高度で詳細な交通データを扱う研究やプロジェクトが注目されている . IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops では AI City Challenge[1] と称し , ITS に関するワークショップを毎年開催している . ここでは , カメラによる車両の識別 , スピード検知 , ナンバープレートの読み込みによるマルチカメラでの車両の再識別などのタスクが用意されており , 国際的に見ても ITS が注目されていることがわかる .

しかし , 現状としてそれらの装置が一般的な道路や交差点に取り付けられていることは少なく , 高速道路や交通量の多い幹線道路に限定し取り付けられていることが多い . これは , 装置自体が高価で大規模化しやすいことが原因であると考えられる . さらに , 日本においてはそれらのデータは , 違反や犯罪 , 渋滞の検知として用いられていることが多いものの , 交通の最適化という観点からデータを活用している事例は少ない . これは , データの質の低さ , ラベルの欠如 , データの利権などが障壁になっていることやそもそも最先端技術を一般道路に活用しようとする働きかけが少ないことが考えられる . しかし , 働きかけは技術の進歩とともに徐々に増えていくものであり事実 , ○○ やトヨタモビリティなど次世代のインフラについて考える機会が増えつつある . したがって , 交差点や道路上の車両の状態を推定するには , IoT デバイスの開発を公開されているデータや安価なエッジセンサで行う必要がある . これらのことから本研究では , 2 次元のカメラ映像から交差点付近にある複数台の車両の追跡をリアルタイムで行うデバイスの開発を目標とした .

こうしたカメラ画像から複数のオブジェクトを検出し追跡するタスクをコンピュータビジョンや機械学習分野では MOT (Multi Object Tracking) と呼ぶ . MOT では , リアルタイムか否かで大きく二つにタスクを分けることができる . リアルタイムではない処理として TNT (TrackletNet Tracker) は , 2 次元のカメラ映像

からオブジェクトを検出し、オブジェクトの軌跡を深層学習を用いて推定していく方法である．[2] [3] TNT は、カメラ映像内で複雑かつ大量のオブジェクトを追跡することに向いているものの、カメラ映像を全て読み込んで処理するため、リアルタイム性はなく、本研究では用いることができない．一方でリアルタイムでの処理として、SORT (Simple Online and Realtime Tracking) がある．[4] [5] これは、検出されたオブジェクトの座標をカルマンフィルタ [6] を用いて、リアルタイムでのオブジェクトの追跡を可能としている．前述で述べた TNT と比べると複雑な MOT を行うことは難しいものの、ある程度規則性を持った車両などの MOT はリアルタイムで処理することが可能である．

第3章

提案手法

本章では，提案手法について説明する．

3.1 提案システム

本研究では、アルゴリズムに SORT，安価なエッジセンサとして Jetson Xavier NX[7]，オブジェクト検出機として YOLOv4[8] を用いて、交差点付近の車両を追跡するシステムを開発した。それらを各信号機に設置し、各センサと各信号機を制御する RaspberryPi4 を接続している。

このシステムを開発する際の問題点として、

- (1) 安価であるが故に十分な性能を発揮することが難しいこと
- (2) 既存のデバイスで得られたデータを活用することが難しいこと
- (3) 環境的に通信の信頼性を担保することが難しいこと

がある。(1) は Jetson Xavier NX を用いて、検証を行う。さらに、SORT は実装が軽量なためリアルタイム処理性能として非常に優れており、本研究ではこれを用いることとした。(2) は The Microsoft Common Objects in Context (MS COCO) dataset[9] を用いて学習された YOLOv4 を使用することで解決する。(3) は金融系で用いられことを目的とした AMQP のプロトコルを使用することで信頼性を担保する。

3.1.1 概要

提案する車両の追跡システムの概要は図に示す。提案するシステムはエッジセンサで車両の追跡を行い、クライアント側で座標データを受け取って、地図上に描画する。エッジセンサ側では、検出器に YOLOv4 を用いて、SORT アルゴリズム [5] を用いて車両の追跡を行う。クライアント側では、ホモグラフィ行列を求め、Google Maps API[10] を用いて地図に車両の位置情報を描画を行う。今回の実験での信号機の制御は、センサから受け取った車両の位置情報から最適な信号制御を行う。

第4章

アルゴリズム

本章では、システムで用いられているアルゴリズムの詳細について述べる。

Algorithm 1 SORT

```

1: 動画の読み込み
2: while do
3:   フレームの読み込み
4:   YOLOv4 でオブジェクト探索
5:   if Object is True then
6:      $X = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$ 
7:      $kalmanfilter(X)$ 
8:   end if
9: end while

```

4.1 アルゴリズム概要

1章で述べたようにエッジセンサ側で SORT , クライアント側で射影変換を用いている .

4.1.1 SORT について

検出器によって検出したバウンディングボックス (BBOX) の座標から , フレーム間の変位を線形モデルを用いて近似させる . 各オブジェクトの状態は次のようにモデル化される .

$$X = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T \quad (1)$$

ここで , u と v はオブジェクトの中心の水平方向と垂直方向のピクセル位置を表し , s と r はそれぞれオブジェクトの BBOX の面積とアスペクト比を表す . ただしアスペクト比は一定とする . 検出された BBOX は , カルマンフィルタのフレームワークを用いて , 速度成分が最適化されるオブジェクトの状態を更新するために用いられる . オブジェクトが検出されていない場合 , 線形モデルを用いて補正せずに , 次の状態を単純に予測する .

4.1.2 射影変換

検出器によって検出した BBOX の底辺の中点 p を射影変換し，地図上に投影する．中点 p は次式のようになる．

$$p = [x, y, 1]^T, P = [X, Y, 1]^T \quad (2)$$

$$P' = [X', Y', W']^T = Hp \quad (3)$$

$$P = \frac{1}{W'} P' \quad (4)$$

ここで， P は地図上の座標を表している． x, y は画像の座標を表しており， H はホモグラフィ行列を表す．

4.1.3 信号制御

本実験では，デモとして最適な信号の制御を行なった．その制御内容は下記の通りである．

- (1) 交差点に接する車道への車両の流入が観測された場合，青信号にする．
- (2) 交差点内で右折待ちで渋滞が発生している場合，対向車側の信号を赤信号にする．

第5章

結果

本章では，提案システムを俯瞰し，実験結果について述べる．

5.1 実験結果と考察

5.1.1 実験概要

本実験では 3 種類のシチュエーションをシミュレーションし、検証を行う。

(実験 1) 初期状態の確認と基本ロジックの確認

(実験 2) ごく少ない交通需要のもとでの信号待ち時間の解消

(実験 3) 右折待ち行列の解消

第6章

謝辞

本研究は、一般財団法人トヨタ・モビリティ基金の支援による「自律分散型信号システム研究開発」の一部として実施した。また、東京大学生産研究所

2021 年 3 月 31 日

安齋 凌介

学外発表

1. 安齋凌介, 岩井将行, ” .

参考文献

- [1] Milind Naphade, Shuo Wang, David C. Anastasiu, Zheng Tang, Ming-Ching Chang, Xiaodong Yang, Liang Zheng, Anuj Sharma, Rama Chellappa, and Pranamesh Chakraborty. The 4th ai city challenge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2020.
- [2] Gaoang Wang, Yizhou Wang, Haotian Zhang, Renshu Gu, and Jenq-Neng Hwang. Exploit the connectivity: Multi-object tracking with trackletnet. In *Proceedings of the 27th ACM International Conference on Multimedia*, pp. 482–490, 2019.
- [3] Zheng Tang, Gaoang Wang, Hao Xiao, Aotian Zheng, and Jenq-Neng Hwang. Single-camera and inter-camera vehicle tracking and 3d speed estimation based on fusion of visual and semantic features. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 108–115, 2018.
- [4] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)*, pp. 3464–3468. IEEE, 2016.
- [5] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and real-time tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)*, pp. 3645–3649. IEEE, 2017.
- [6] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

- [7] NVIDIA. Jetson xavier nx, 2021. <https://www.nvidia.com/ja-jp/autonomous-machines/embedded-systems/jetson-xavier-nx/>.
- [8] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.
- [9] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pp. 740–755. Springer, 2014.
- [10] Google. Google maps platform. <https://developers.google.com/maps/?hl=ja>.

付録

In the Appendix, attach a paper was presented at the Entertainment Computing 2015.