

# 運転挙動及び車載カメラ画像を用いた潜在リスク交差点検知手法

豊田 正史<sup>†</sup> 横山 大作<sup>†</sup> 小宮山純平<sup>†</sup> 伊藤 正彦<sup>†</sup>

<sup>†</sup> 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1  
E-mail: †{toyoda,yokoyama,jkomiya,imash}@tkl.iis.u-tokyo.ac.jp

あらまし 交通事故の件数は減少傾向にあるものの業務用車両の事故件数に関しては横這いが続いており、運送事業者等においては事故防止が依然として重要な課題である。本論文では、運転者や通行者が注意を払うべき潜在リスクを持つ交差点を大規模なドライブレコーダデータから得られる運転挙動データ及び車載カメラ画像から検知する手法を提案する。本手法では、大規模な運転操作データに加え、運転が行われた道路の詳細な形状および車載カメラ画像から抽出した車両や歩行者の出現数等を考慮して実際に事故が発生した交差点と特徴が類似する交差点を潜在リスクとして検知する。数千人規模のドライブレコーダデータを用いた実験を行いその有効性を示す。

キーワード 交通, ドライブレコーダ, 潜在リスク

## 1. はじめに

内閣府の平成 29 年版交通安全白書 [1] によると、交通事故発生件数は近年減少を続け平成 27 年には年間約 53 万 7 千件となったものの、自動車保有台数 1 万台当たりの死傷者数は、平成 3 年から横ばいになり近年では漸減傾向が続いており、依然として深刻な社会的課題である。業務用車両においては一定以上の大きさの車両についてはドライブレコーダの装備が義務付けられているものの事業者では、いまだ事故や死傷者数を減少させる取り組みが盛んであり、運転データ等に基づく精密な交通安全支援は今後も重要となる。

業務用ドライブレコーダの市場は現在も拡大を続けているため交通安全に資するデータが大規模に獲得できる状況が実現している。ドライブレコーダはもともと事故の瞬間を記録するものであったが、近年では高機能化が進んでおり、記録できる情報の量も種類も増えてきている。事故の記録だけでなく、ブレーキやハンドル等の運転操作を常時記録できるようになってきており、大規模な運転操作データが蓄積されている。

本論文では、大規模な業務用ドライブレコーダデータを用いて、道路上での運転操作の特徴に関して詳細に分析し、さらに運転が行われた道路や、車載カメラで撮影された画像などの異種データを統合して潜在的リスクのある道路を検知する手法を提案する。ここでは交差点の事故に着目し、実際に交通事故が発生した交差点における特徴を学習し、特徴の類似した交差点を潜在リスクを持つものとして検知する。1 つの事故の背後には、多数のヒヤリハット事例が隠れており、運転操作の分析から交通事故に関するヒヤリハットを抽出し、潜在リスクとして示すことができれば事故防止への活用が期待できる。例えば、運送・運輸事業者におけるドライバーの教育や、自治体におけるヒヤリハット地図の提供、道路修繕の優先順位付けなどへの利用が考えられる。現在、多くの自治体では、住民にアンケートを取るなどして道路の危険箇所を示した「ヒヤリハット地図」を作成しているが、主に人手で作成されているためカバーできる範囲や正確性に問題がある。また警視庁では実際に事故が発

生した箇所を「交通事故発生マップ」として公開している<sup>(注1)</sup>が、実際に事故が発生した場所のみを注意するだけでは不十分である。

我々は、ドライブレコーダのメーカーと協力し、運送事業者やバス事業者の大規模なドライブレコーダデータ、道路の数値地図等の異種データを用いた交差点の潜在的リスクの推定手法の研究を行っている。これまでに、バス事業者のドライブレコーダデータ、道路地図、および雨量情報を用いたリスク推定手法を提案してきたが [2]、本論文では、車載カメラ画像から抽出した車両や人などの情報を用いたリスク推定手法を試みる。実験の結果、交通事故リスクに寄与する特徴量に関するいくつかの知見が得られたためこれを報告する。

本論文の構成は以下の通りである。2. 節では関連研究を述べる。3. 節では研究に用いたデータセットの詳細を説明する。4. 節ではドライブレコーダデータ、道路地図、および車載カメラ画像を考慮した、データ統合手法ならびに潜在リスク交差点検知手法を提案する。大規模な業務用ドライブレコーダデータを用いた実験結果を 5. 節で述べる。6. 節でまとめと今後の課題について述べる。

## 2. 関連研究

カーナビなどから得られた危険運転が行われた場所を提供するサービスとして本田技研工業株式会社のインターナビ<sup>(注2)</sup>や東京都トラック協会の G Map<sup>(注3)</sup>がある。実装の詳細は不明であるが、これらのサービスでは急ブレーキや急ハンドルの発生等、単純な指標しか用いておらず、本研究で用いているような詳細な運転操作の属性については考慮されていないようである。

ドライブレコーダデータから得られる詳細な運転操作データを分析した研究としては、[3-5] などがあるが、いずれもドライバーの分類を目的としており、道路のリスクに関する分析は

(注1): <http://www3.wagamachi-guide.com/jikomap/>

(注2): <http://www.honda.co.jp/internavi/>

(注3): <http://gmap.dgis.jp/dc/ngt.html>

行われていない。[3,4]では少数のドライバーを詳細に分析することを試みており実験から得られるデータを道路リスクの分析に用いることは難しい。我々の研究グループは[5]において本研究と同種のドライブレコーダから得られた大規模なデータを用いてドライバーのリスクを推定する手法を研究している。このような1,000人規模のデータになると道路の網羅率も高くなり、道路リスクを分析することが可能となる。

道路の事故リスクに関する研究では、これまで主に道路の形状や環境および交通量や事故データを用いて分析が行われてきた[6-8]。これらの研究では、交差点等における事故の発生率を様々な道路特徴から推定する関数を Safety Performance Function (SPF) と称し、道路の幅やレーン数、曲率などの特徴や交通量、事故発生データなどから事故率を回帰等によって求め、事故の発生原因などを分析している。しかし、交通データに関しては、交通量や平均速度など基本的な統計量のみを用いており、詳細な運転操作データを用いたものは見当たらない。

交通事故発生データの視覚的な分析を行った研究としては、[9-11]などがあるが、主に事故の統計的な分析や閲覧を目的としたもので、道路の潜在的なリスクを把握しようとする試みはなされていない。我々は、ブレーキ操作やハンドル操作における車両の加速度や躍度といった詳細な情報を可視化し、要注意道路を視覚的に探索するシステムを開発してきたが[12]。手動による探索には限界があり、潜在リスク箇所を自動的に検知する手法による支援が必要不可欠である。

### 3. データセット

本節では、本研究で用いたデータセットの詳細について述べる。本研究では業務用ドライブレコーダを設置している運送事業者から提供を受けたドライブレコーダデータおよび交通事故データを用いている。さらに道路状況、運転状況を考慮するため道路数値地図データ、および車載カメラ画像データを用いている。

#### 3.1 ドライブレコーダデータ

ドライブレコーダのメーカーである株式会社データテックの協力の元、顧客である運送事業者から実際に運行されている配送車両のドライブレコーダの提供を受けた。本データは、約1500人のドライバーの約3年分の走行データを含む大規模なものである。本データは、GPSから取得した0.5秒毎の車両の軌跡データと、ブレーキやハンドル操作といった操作の際に記録される運転操作データを用いている。

本実験で用いたドライブレコーダは、危険運転時に限らず全ての運転操作を記録している。各種運転操作を走行中に検知し、操作毎に操作が行われた時の時間、位置、速度や加速度等の代表値を記録する。例えば、ブレーキ操作が検知されると、そのブレーキ操作が行われていた時間帯中の最大速度、最大前後加速度、最大前後加速度変化(ジャーク)の値がブレーキ操作の代表値として記録される。本実験では、ブレーキ操作とハンドルの2種類の操作記録を用いている。記録されているデータの属性の詳細を表1に示す。

#### 3.2 数値地図データ

運転操作の行われた交差点の道路状況を取得するために、道路の数値地図データを用いる。本実験では、住友電工の拡張版全国デジタル道路地図DBを用いている。本地図データには、交差点(ノード)の緯度経度情報、交差点の名前や信号の有無等の属性情報、および交差点に接続している道路(リンク)の属性情報が収録されている。リンク属性としては、道幅、交通制限などの情報が利用可能である。拡張版全国デジタル道路地図DBにおいては、各リンクに4段階(0~3m, 3m~5.5m, 5.5m~13m, 13m~)の道幅情報、レーン数が付与されている。

本手法では、今回走行データが取得できている運送事業者の営業地域における交差点を上記数値地図より抽出し、各交差点の特徴量として、信号の有無、接続しているリンクの本数、道幅の段階毎のリンク本数分布、レーン数毎のリンク本数分布、および交差点に接続しているリンク間の最小角度を用いる。

#### 3.3 事故データ

警視庁では実際に事故が発生した箇所を「交通事故発生マップ」として公開している(注4)。本実験においては、上述した各交差点において、3年間の間に起きた車両関連の事故件数をカウントし、1回以上交通人身事故が発生した交差点を正例、1回も事故が発生していない交差点を負例として用いる。

#### 3.4 車載カメラ画像データ

今回データを取得したドライブレコーダの一部は車載カメラを搭載しており、30秒ごとに撮影した静止画が保存されている。今回の実験では、車載カメラで3年間の間に撮影された800万枚の画像を用いる。まず、各交差点について中心から30m以内で撮影された画像を抽出して、交差点毎の画像のセットを作成する。各画像からFaster R-CNN[13]という一般物体認識手法を用いて、車、人、バス、自転車、バイクの5種類の物体を認識し、認識された物体の写真1枚あたりの平均出現数を特徴量として用いる。

## 4. 潜在リスク交差点検知手法

本節ではまず、ドライブレコーダデータに含まれる各運転操作に対して、数値地図データを付与するための方法について述べ、その後、統合したデータを用いて潜在リスク交差点を検知するための手法について述べる。

#### 4.1 運転操作データへの数値地図データの付与

提案手法では、交差点のリスクに関する特徴を、その交差点において行われた運転操作の集合、交差点の特徴、および周辺で撮影された写真を用いて表現する。このため本手法では、運転操作を基本データとし、各運転操作に対して、その運転が行われた時の道路の特徴を紐づける統合化処理を行う。その詳細を以下に述べる。

(1) 各運転操作の前後の時点での道路を特定するため、操作時の座標、および前後3秒の時点における座標をGPSの軌跡データから取得する。

(2) 数値地図データについて、座標から近傍の道路リンク

(注4): <http://www3.wagamachi-guide.com/jikomap/>

表 1 運転操作データの詳細

運転操作	共通属性	操作特有の属性
ブレーキ操作	時刻, 緯度, 経度	最大速度, 最大前後加速度, 最大前後加速度変化
ハンドル操作	時刻, 緯度, 経度	最大速度, 最大角加速度, 最大角加速度変化, 最大左右加速度

を取得するインデックスを作成し、これを用いて上記 3 地点を最近傍の道路リンクへ射影する。この 3 地点に対応する緯度経度、リンク ID、道幅、交通規制などの必要情報を当該運転操作に付加する。

この処理により、各運転操作から道路の情報を参照することが可能となる。データの冗長性は増すことになるが、実験の過程において運転操作のスキャンおよび運転状況データの参照を繰り返すことになるため、統合化を行う方が効率的である。運転状況が付加された本運転操作データは、道路リスクの分析のみならずドライバーの分析にも利用可能であり汎用性も高い。

#### 4.2 交差点における運転操作に関する特徴量

上記の統合化データを用いて、交差点のリスク検出のための特徴量を設計した。まず、数値地図を用いて運送事業者の営業エリア内の交差点をすべて抽出し、各交差点に関して 30m 以内で行われた全ての運転操作を抽出する。さらに、操作数が一定以上ある交差点のみを実験対象として抽出する。この運転操作の集合を用いて交差点の特徴を表現することになる。本手法で用いる特徴量の詳細を以下に述べる。

**操作数** 交通量の多い交差点は事故の発生する可能性が高いことを反映するため、交差点で行われた操作回数を疑似的な交通量として用いる。ブレーキ操作およびハンドル操作の 2 種類の回数を使用する。

**操作属性値の分布** 交差点で行われた操作の属性値の分布を特徴量として用いる。ここでは、各属性値（最大速度、最大加速度等）の平均値および標準偏差を使用する。

**異常操作数** 急ブレーキ、急ハンドル等の異常操作が行われた回数が多いと事故が発生する可能性が高くなるという仮説を反映するため、異常操作を検出しその数を特徴量として用いる。最大速度と最大加速度、最大加速度と最大加速度変化など属性値の間にはある程度の相関関係が認められるため、マハラノビス距離を用いた異常値検出手法を使用する。特徴量としては、マハラノビス距離が閾値以上となる操作の数をを用いる。今回の実験では複数の閾値 (2,4,6,8) を用い、各閾値を超える操作の数を特徴量として用いた。

**道幅毎の操作数** 交差点では異なる道幅の道路が交わることがあり、太い道から細い道へ右左折で進入したり、交通量の多い太い道へ細い道から合流したりする場合などに事故リスクが増加する。こうした運転状況を考慮するため、操作 3 秒前の道幅と 3 秒後の道幅のペアをパターンとして用い、パターンごとに操作集合を分割し、各パターンの操作数を特徴量として用いる。4 段階 (0m ~ 3m, 3m ~ 5.5m, 5.5m ~ 13m, 13m ~ ) の道幅の組み合わせ (計 16 パターン) を使用する。

**道幅毎の操作属性値の分布** 道幅のパターンで分割した操作集合それぞれに対して、属性値の平均および標準偏差を特徴量として用いる。

**道幅毎の異常操作数** 道幅のパターンで分割した操作集合それぞれに対して、上記の方法で異常操作を検出し、同じ閾値のセットを用いた異常操作数を特徴量として用いる。

#### 4.3 交差点の道路形状に関する特徴量

各交差点の道路形状に関する特徴量を数値地図のデータから抽出して用いる。以下に述べるように、信号の有無、接続しているリンクの本数、道幅の段階毎のリンク本数分布、レーン数毎のリンク本数分布、および交差点に接続しているリンク間の最小角度を用いる。

**信号の有無** 各交差点における信号の有無を 1/0 で表す。

**リンクの本数** 交差点に接続している道路リンクの本数を用いる。ただし、中央分離帯等がある交差点の場合、1 つの交差点が上下車線毎に設定された複数のノードで表現される場合があるため、短い (20m 程度) リンクで接続されたノードがある場合にはそれらを併合し、交差点外へ伸びるリンクの本数を特徴量として用いる。

**道幅毎のリンク本数分布** 交差点に接続している道路リンクの道幅の分布を用いる。今回用いた数値地図では上述のように 4 段階の道幅があり、各道幅のリンクが何本接続しているかを 4 要素のベクトルで表現する。

**レーン数毎のリンク本数分布** 交差点に接続している道路リンクのレーン数の分布を用いる。今回用いた道路地図には、各リンクのレーン数が登録されており、レーン本数毎に何本のリンクが接続しているかをベクトルで表現する。

#### 4.4 交差点の周辺で撮影した車載カメラ画像に関する特徴量

各交差点の中心から 30m 以内で撮影された画像をデータセットから抽出して、一般物体認識手法を用いて認識した車や人などの物体の数をを用いる。具体的には各画像から Faster R-CNN [13] 手法を用いて、車、人、バス、自転車、バイクの 5 種類の物体を認識し、認識された物体の写真 1 枚あたりの平均出現数を 5 次元のベクトルとして表現し特徴量として用いる。

#### 4.5 潜在リスク交差点の検知

上記で設計した特徴量を用い、交差点における事故発生データを教師データとした教師あり機械学習により、潜在リスク交差点を検知する。交差点を事故発生のありなしで分類する分類器を用いる場合は、事故未発生であるにも関わらず事故ありと判定された交差点を潜在リスク交差点と考えることができる。また、交差点の事故回数に対する回帰を用いれば、事故の確率の高い順に交差点を順序付けすることで、潜在リスク交差点のランキングとして用いることができる。

## 5. 実験結果

今回の実験では、事故が発生した交差点のラベルを事故回数に関わらず正例とし、事故の発生していない交差点を負例とし

表 2 潜在リスク交差点検知実験の精度

特徴量	ROC AUC
道路地図	0.797
運転操作	0.728
車載カメラ画像	0.625
道路地図 + 運転操作	0.801
運転操作 + 車載カメラ画像	0.735
道路地図 + 車載カメラ画像	0.801
全特徴	<b>0.803</b>

て、分類を行うことで潜在リスク交差点の検知を試みた。まず、150 回以上のブレーキ操作が行われた交差点を抽出し分析の対象とした。対象交差点は 16,405 あり、そのうち事故が発生した交差点の数は 7,559 箇所であった。

道路地図から得られた交差点の特徴、運転操作に関する特徴、および車載カメラ画像から得られた物体数の特徴量の組み合わせに対して、XGBoost [14] と呼ばれるツリーブースティング分類器を用い、ROC AUC が最大となるようにパラメタのチューニングを行った。本実験では、交差点を 5 セットに分割し、そのうち 4 セットで 4-fold クロスバリデーションによりパラメタチューニングを行い、残り 1 セットでテストを行うことを、5 回繰り返し ROC AUC の平均により評価を行った。

実験結果を表 2 に示す。道路地図、運転操作、車載カメラ画像の 3 種類の特長量をそれぞれ単独で用いた場合、道路地図から得た特徴を用いた時に最も良い ROC AUC となっている。特徴量の組み合わせでは全ての特徴量を用いた時に僅かではあるが最大の ROC AUC を示している。全特徴量から一種類のみ除いた場合には、道路地図の特徴を除いた場合が一番 ROC AUC の減少が大きく、道路の形状がリスクに大きな影響を持っていることが分かる。

## 6. まとめと今後の課題

本論文では、運転者や通行者が注意を払うべき潜在リスクを持つ交差点を大規模なドライブレコーダデータに加え、運転が行われた道路の形状や車載カメラ画像を考慮して検知する手法を提案した。運送事業者の大規模なドライブレコーダデータ、道路の数値地図、車載カメラ画像という異種データを統合した特徴量を設計するためのデータ統合化方法を示し、操作数やその属性分布、異常操作数などを用いた特徴量を用いた、潜在リスク交差点の検知手法を設計し、実験を行った。実験結果より、提案したすべての特徴を用いた際に最大 0.8 という高い ROC AUC が得られた。

- [1] 内閣府. 平成 29 年版交通安全白書. 2017.
- [2] 豊田正史, 横山大作, 伊藤正彦. 運転状況を考慮したドライブレコーダデータからの潜在リスク交差点検知手法. 第 9 回データ工学と情報マネジメントに関するフォーラム (DEIM2017), A5-1, 2017.
- [3] Y. L. Murphey, R. Milton, and L. Kiliaris. Driver's style classification using jerk analysis. In *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, pp. 23–28, March 2009.
- [4] D. A. Johnson and M. M. Trivedi. Driving style recognition using a smartphone as a sensor platform. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1609–1615, Oct 2011.
- [5] D. Yokoyama and M. Toyoda. Do drivers' behaviors reflect their past driving histories? - large scale examination of vehicle recorder data. In *2016 IEEE International Congress on Big Data (BigData Congress)*, pp. 361–368, June 2016.
- [6] Ketong Wang, Jenna K. Simandl, Michael D. Porter, Andrew J. Graettinger, and Randy K. Smith. How the choice of safety performance function affects the identification of important crash prediction variables. *Accident Analysis and Prevention*, Vol. 88, pp. 1 – 8, 2016.
- [7] Francesca Russo, Mariarosaria Busiello, and Gianluca Dell'Acqua. Safety performance functions for crash severity on undivided rural roads. *Accident Analysis and Prevention*, Vol. 93, pp. 75 – 91, 2016.
- [8] Alfonso Montella and Lella Liana Imbriani. Safety performance functions incorporating design consistency variables. *Accident Analysis and Prevention*, Vol. 74, pp. 133 – 144, 2015.
- [9] Anna Fredrikson, Chris North, Catherine Plaisant, and Ben Shneiderman. Temporal, geographical and categorical aggregations viewed through coordinated displays: A case study with highway incident data. In *Proceedings of the 1999 Workshop on New Paradigms in Information Visualization and Manipulation in Conjunction with the Eighth ACM International Conference on Information and Knowledge Management, NPIVM '99*, pp. 26–34, New York, NY, USA, 1999. ACM.
- [10] M. L. Pack, K. Wongsuphasawat, M. VanDaniker, and D. Filippova. Ice-visual analytics for transportation incident datasets. In *2009 IEEE International Conference on Information Reuse Integration*, pp. 200–205, Aug 2009.
- [11] A. Anwar, T. Nagel, and C. Ratti. Traffic origins: A simple visualization technique to support traffic incident analysis. In *2014 IEEE Pacific Visualization Symposium*, pp. 316–319, March 2014.
- [12] M. Itoh, D. Yokoyama, M. Toyoda, and M. Kitsuregawa. Visual interface for exploring caution spots from vehicle recorder big data. In *2015 IEEE International Conference on Big Data (Big Data)*, pp. 776–784, Oct 2015.
- [13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems 28*.
- [14] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, pp. 785–794, 2016.