

# ドライブレコーダデータと道路特徴を併用した 潜在リスク箇所の判別に向けて

小矢島 諒† 豊田 正史†† 横山 大作†† 伊藤 正彦††

† 東京大学大学院情報理工学系研究科 〒113-0033 東京都文京区本郷 7-3-1 東京大学 情報理工学系研究科

†† 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1 東京大学 生産技術研究所

E-mail: †{koyajima, toyoda, yokoyama, imash}@tkl.iis.u-tokyo.ac.jp

**あらまし** 本研究では、ドライブレコーダから得られた大量の車両の運転操作の情報と道路特徴データを用いることで、道路のリスクを把握することを試みる。まず、運転操作の発生分布を地図上に可視化し、インタラクティブに探索するインターフェースを提案する。これにより、操作が集中している箇所を効率的に探索し、個々の操作集中箇所について詳細な分析を行うことを可能とした。このインターフェースを用いて事故の有無と運転操作の分布の関係を分析した。また、リスクとなり得る道路特徴のラベル付けを行い、事故の有無と道路特徴の有無の関係を分析も行った。これらの分析結果を基に、交通事故が発生した交差点を判別することでリスクを持つ交差点の抽出を試みた。この結果、運転操作に加えて道路特徴も特徴量として用いることで判別性能が向上することを確認した。最後にこの判別結果を基に、事故は未発生にも関わらず、発生したと判別された交差点についてリスクの分析を行った。

**キーワード** ドライブレコーダ, 情報可視化, 可視化インターフェース, データマイニング

## 1. はじめに

警視庁の統計<sup>(注1)</sup>によると、平成26年中の交通事故件数は約57万件に上り、ここ数年は年々減少しているとはいえ、未だ大きな社会問題として存在している。ハインリッヒの法則という、労働災害における経験則では、「1つの事故の背後には、29の軽微な事故と300のヒヤリハットが隠れている」とされている[1]。これを交通事故に当てはめると、交通事故の背後には多くのヒヤリハットが発生していると考えられる。この法則には「ヒヤリハットを防ぐことで事故も防げる」ことも示されており、ヒヤリハットが起きるようなリスクのある道路に対策を施すことができれば、事故防止に大いに役立つことが期待できる。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<sup>(注2)</sup>。しかしながら、こうした地図は人手で作成されているため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3)</sup>が、危険箇所の網羅性は非常に悪い。そこで、正確で網羅性のあるヒヤリハット地図が求められている。この問題に対し、ドライブレコーダデータを利用する動きがある。事故を記録することを目的としたドライブレコーダの普及が進んでおり、GFK社<sup>(注4)</sup>の調査によると、ドライブレコーダの国内販売台数は、2013年度におよそ30万台、2014年度には前年比1.5倍の

43万台に達した。国内の累計販売台数はおよそ100万台とみられ、これからも普及していくと考えられる。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。そのため、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。こうした背景から、ドライブレコーダから得られる情報を大量に集めて道路状況を分析し、活用しようという取り組みが増えてきている。

本研究では、ドライブレコーダデータを用いて道路上での運転操作を詳細に分析し、個々の道路特徴を併せて道路のリスクを把握することを目指す。まず、運転操作を地図上に可視化することで、道路上での運転操作分布をインタラクティブに分析するインターフェースを提案する。ブレーキ操作やハンドル操作に関する速度・加速度等の詳細なデータを用いることで、運転に際して注意を要する道路の探索や、その道路でどのような運転操作の傾向が存在するかを分析することができる。また、Google ストリートビュー<sup>(注5)</sup>を基に、交通事故のリスクとなりえる道路特徴のラベル付けをインターフェース上で行うこともできる。このインターフェースを用いて、事故の有無と運転操作の分布、道路特徴の有無との関係を分析した。さらに、事故の発生した交差点の判別実験を行った。この実験では、特徴量として、運転操作に加え、先にラベル付けを行った道路特徴を用いることで判別性能がどのように変化するかを確認した。最後に、この事故判別結果を基に、事故は未発生だが事故が発生したと判別された交差点について、どのようなリスクがあるかを分析した。

(注1) : <http://www.npa.go.jp/toukei/koutuu48/before/hasseijokyo/PDF/H26hasseijokyo.pdf>

(注2) : <http://www2.wagamachi-guide.com/shinagawa/index.asp?dtp=1>

(注3) : <http://www3.wagamachi-guide.com/jikomap/>

(注4) : <http://www.gfkjpn.co.jp/jp/news-and-events/press-room/press-releases/Documents/150227%20driving%20recorders.pdf>

(注5) : <https://www.google.com/intl/ja/maps/streetview/>

表 1 ドライブレコーダデータ

| 運転操作   | レコード数     | 属性                                                     |
|--------|-----------|--------------------------------------------------------|
| ブレーキ操作 | 1,242,783 | 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 | 1,583,476 | 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                        |
| 右左折操作  | 657,066   | 進入前最小前後加速度, 進入速度, 最大速度, 最大合成G, 方位角速度ゆらぎ, 終了時速度, 終了時合成G |
| ブレーキ操作 | 1,526,510 | 最大速度, 最大前後加速度, 停止時間                                    |

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を紹介する。4. 節ではドライブレコーダデータと道路特徴データを用いた分析・可視化システムについて述べ、実際にシステムを用いて運転操作と道路特徴を分析した結果を5. 節で述べる。この分析結果を基に事故の発生した交差点を判別する実験を行った結果と、判別の結果事故が発生していないが事故が発生したと判別された、潜在リスクのある交差点についての考察を6. 節で述べる。最後に7. 節でまとめと今後の課題について述べる。

## 2. 関連研究

危険な運転が行われた場所を表示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6)</sup>や東京都トラック協会のG Map<sup>(注7)</sup>がある。しかしながら、これらのサービスでは急ブレーキや急ハンドル等の単純な指標しか用いておらず、危険箇所の網羅性が不十分である。ドライブレコーダデータのうち、GPS データを用いて車両の軌跡を分析した研究は様々な国で数多く行われており、アメリカのタクシーのデータを分析したものとして[2,3,14]などがあり、スウェーデンのタクシーデータを分析したものとしては[13]がある。Ferreira らは街間の移動時間を可視化する手法を提案している。また中国のタクシーのデータを分析したものとしては、[4-12]などがある。Wang らは、ドライブレコーダのGPS と速度の情報に加え、街の道路マップを使用した。ドライブレコーダのデータから各車両の軌跡を作成し、その軌跡の座標と道路マップの座標をマッチさせ、実際の道路に各軌跡をマップした。そうして、道路ごとにマップされた軌跡データから車両の速度を計算し、ある閾値以下の道路では渋滞が起きていると定義した。こうして渋滞を検出した後は、渋滞の伝播を調べるため、時系列ヒートマップを作成し、各道路上では何時に渋滞が起きているかを可視化した。さらに、先の渋滞図と比べることで、図の例ではA点で起きた渋滞が、その先の道路であるB、C点に伝播しているのがわかる。ミラノの車両の流れや軌跡を分析したものとしては[15]がある。Andrienko らは、車両の軌跡をまとめて可視化することで車両の流れを視覚的に分析した。これは、街の交通マネジメントを目的としたものであった。Chu らは、タクシーの軌跡を文書に変換し、LDA を使うことで軌跡のパターンを発見した[10]。Ferreira らはニューヨークのタクシーのGPS データを用いて街間の移動時間を可視化する手法を提案している[3]。これらの研究は主に車両のGPS データを用いており、加速度や躍度のような詳細な情報の分析は行われていない。

それらの詳細な情報を分析した研究としては、[16,18,19]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、道路の分析は行っていない。交通事故の分析を行った研究としては、[20-22]などがあるが、交通事故が未だ起きていないがリスクがあると考えられる場所を把握しようとする試みはなされていない。車両の加速度や躍度といった詳細な情報や、ブレーキ操作やハンドル操作などの運転操作を分析した研究としては、Itoh らの研究がある[23]。彼らは車両の加速度や躍度といった情報や、ハンドル操作やブレーキ操作などの運転操作から危険領域を探索する手法を提案した。しかし、道路特徴による運転操作の傾向の分析や、運転操作を用いて道路の潜在リスクを判別しようとする試みは行われていない。

## 3. データセット

本稿では、ドライブレコーダデータ、数値地図データ、交通事故データの3つのデータセットを用いている。

### 3.1 ドライブレコーダデータ

ドライブレコーダデータとしては、株式会社データテックの協力の下で、運送会社のドライバーによる実データを用いた。収集期間は2014年7月から2015年7月までの約1年である。利用した区域は文京区を中心とした東京都近辺である。本稿では、このデータのうち、0.5秒ごとに記録される車両の軌跡データと、ハンドルやブレーキといった運転操作の際に記録される操作ログを用いている。ドライブレコーダデータの内容を表1にまとめた。運転操作には4つの種類がある。ハンドル操作とブレーキ操作は、それぞれの操作が行われると記録される。右左折操作は、交差点で右左折を行った時に記録され、ハンドル操作とは独立している。停止操作は車両が停止する際に記録される。

### 3.2 数値地図データ

数値地図データとしては、東京大学空間情報科学研究センター研究用空間データ基盤より株式会社ゼンリンによるZmap TOWN II<sup>(注8)</sup>の提供を受けて利用している。データはノードとリンクの形で分かれており、交差点等はノードとなり、これらを繋ぐ道路がリンクとして記録されている。本稿で用いた数値地図データを表にまとめた。

表 2 数値地図データ

|     | 属性                                              |
|-----|-------------------------------------------------|
| ノード | 交差点かどうか<br>信号の有無                                |
| リンク | 道路の幅 (0~3m, 3m~5.5m, 5.5m~13m, 13m~)<br>一方通行の有無 |

(注6) : <http://www.honda.co.jp/internavi/>

(注7) : <http://gmap.dgis.jp/dc/ngt.html>

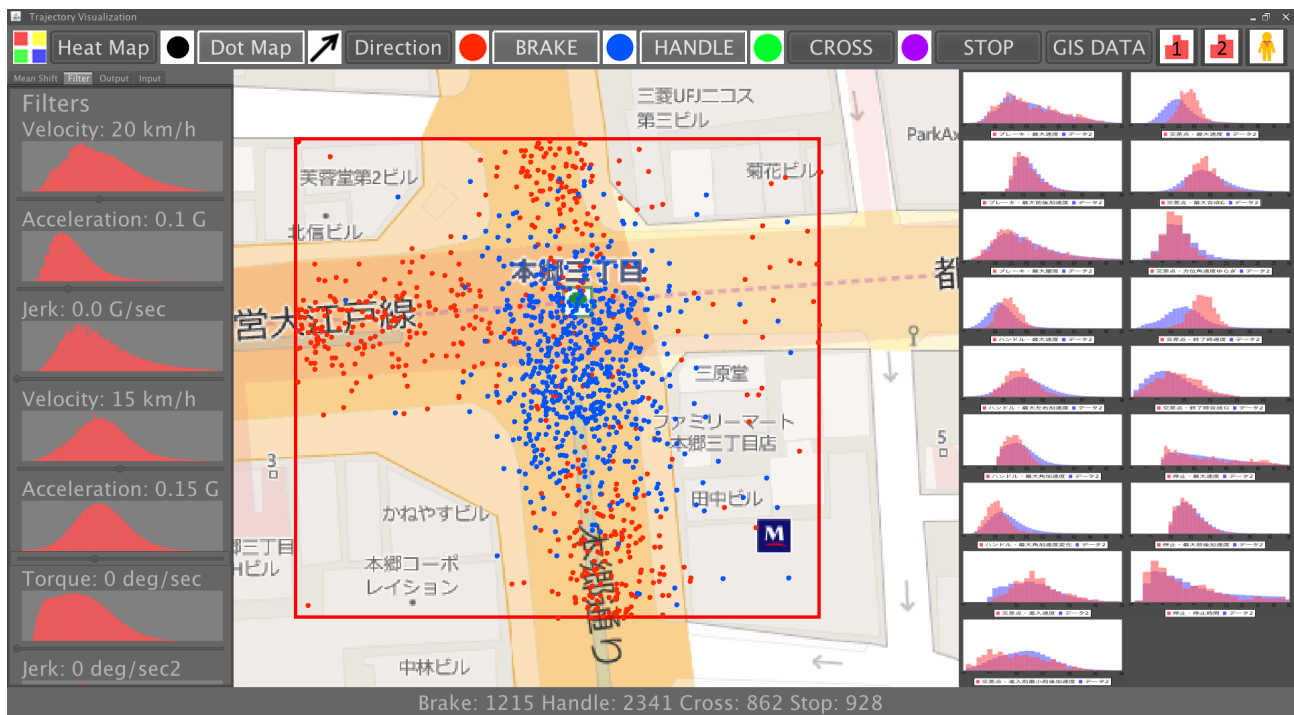


図 1 分析・可視化システム

### 3.3 交通事故データ

交通事故データとしては、警視庁のホームページで公開されている、平成 25 年度から平成 27 年度の上半期までの 2 年半分の東京都での交通事故情報を用いた。内容としては、交通事故 1 件に対してその位置と記録された年度が含まれている。

## 4. 分析・可視化システム

本稿ではドライブレコーダーデータと道路特徴データを地図上に可視化し、運転操作を探索するためのインターフェースを提案する。本インターフェースでは、指定した条件を満たす運転操作が集中している箇所を見つけ、その箇所での運転操作の詳細な分析を行うことを目的としている。運転操作はブレーキ、ハンドル、右左折、停止の 4 種類があり、それぞれを同時に別々にも探索することができる。分析・可視化システムの全体図を図 1 に示す。また、潜在リスク箇所を探索するワークフローを図 2 に示す。探索の流れとしては、まず操作発生頻度を表すヒートマップを用いて、運転操作が集中している箇所を広域的に探す。この際、フィルタリングを行い、対象を絞って探索することも可能である。詳細に分析したい道路がある場合、操作発生地点のドットマップを用いることで、道路のどこで運転操作が行われたかを知ることができる。各運転操作には操作前後 3 秒の車両の位置情報が紐付けてあるため、操作前後の車両の軌跡も合わせて表示することで、車両がどこからどこへ向かう時に操作が行われたかが分かる。さらに、ある道路で記録された運転操作全体の傾向を知りたい場合、その道路を地図上で選択することで、その範囲で記録された運転操作の分布をヒストグラムで表示する。また、道路を複数選択することで、そ

れらの範囲で記録された運転操作の分布を比較し、傾向の違いを分析できる。加えて、道路状況を詳しく知りたい場合はインターフェース上のボタンを押すことで素早くストリートビューを見ることができ、実際の道路状況を把握可能である。

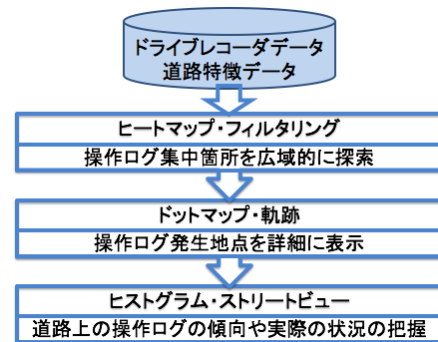
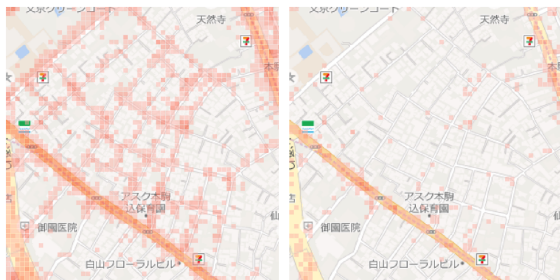


図 2 分析・可視化システムのワークフロー

まず、運転操作が集中している箇所を広域的に探索するためにヒートマップを用いる (図 3(a))。ヒートマップは 10m 四方の矩形の集合で構成されている。運転操作ごとの探索を可能にするため、種類ごとに異なる色を割り当て別々にヒートマップを作成した。また、各運転操作を同時に探索できるように、各ヒートマップの階級には透明度を割り当てる。透明度は矩形内の運転操作の数に応じて 5 段階で変わり、ログの数が多いほど透明度が低くなるように表示する。

潜在リスクの存在するような道路では、危険度の高い運転操作が多く記録されている可能性がある。そのため、本インターフェースでは表示する運転操作のフィルタリングが可能である。図 1 のように、運転操作の閾値をスライダーのつまみを動かす

(注8) : <http://www.zenrin.co.jp/product/gis/zmap/zmaptown.html>



(a)フィルタリング前

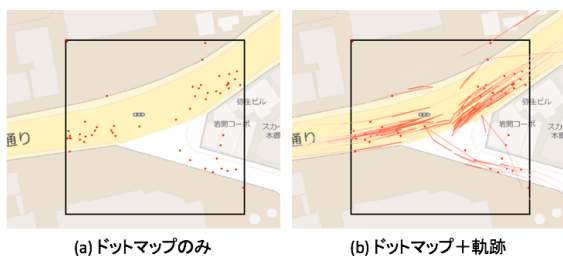
(b)フィルタリング後

図3 操作発生頻度のヒートマップ

ことで、表示するログの数を対話的に変更することができる。この際、各運転操作の各属性ごとに値のヒストグラムをスライダー上部に表示しているため、現在の閾値がその属性の値全体のどの位置にいるかを知ることができる。

ヒートマップを表示中にフィルタリングが行われると、ヒートマップの透明度はフィルタリング後の運転操作のみを用いて再度5段階で計算される。これにより、図3(b)のようになり、フィルタリング後のヒートマップの透明度が偏ってしまうことを防いでいる。

各運転操作が記録された実際の場所はドットマップで表示できる(図4(a))。これにより、道路のどの位置で運転操作が行われたのかを見ることができる。さらに、各運転操作が記録された前後3秒の車両の位置も表示が可能である(図4(b))。この7秒分の車両の軌跡は、秒ごとに直線の濃度を段階的に変えている。操作3秒前から2秒前を結んだ線分の濃度を最も薄く、また操作2秒後から3秒後までの線分を最も濃く設定することで、操作がどの方向からどの方向へ移動している際に記録されたかが分かる。



(a)ドットマップのみ

(b)ドットマップ+軌跡

図4 操作発生地点のドットマップとその前後3秒の車両の位置を表示した様子

ある道路上で記録された運転操作全体を分析するために、本インターフェースでは運転操作の分布の傾向を見ることができる。図5のように地図上で分析したい範囲を矩形で選択することで、その範囲で記録された全ての運転操作のヒストグラムを、各操作の各属性ごとに表示する。これらヒストグラムはそれぞれログの数で正規化しているため、どれくらいの割合で分布しているかを知ることができる。また、この矩形は2つまで選択することができ、2つの場所で記録された運転操作の分布を重ねて表示することで、分布の傾向を比較することができる。さらに、この比較は矩形同士だけでなく、全ての運転操作との比

較や、特定の道路特徴を持つ道路全ての運転操作とも比較することができるため、様々な角度から分布の傾向の違いを分析することができる。

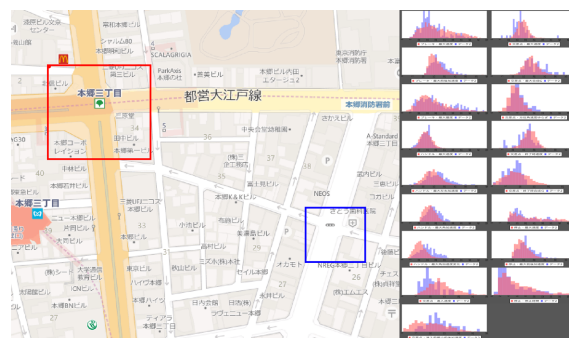


図5 選択した矩形における運転操作の分布を比較する様子

道路地図上では、実際の道路状況を把握することができない。そこで、地図上で任意の座標を選択し、ストリートビューを見ることができるようにした。図1に設置されたボタンを押すだけですぐにストリートビューに飛ぶことができ、素早い分析作業が期待できる。これにより、例えば道路のどこに電柱が立っているか、あるいは路側帯に車が止まっているかなど、地図では知り得ない道路特徴を把握することができる。さらに、これによって発見した道路特徴は、図1左側にある「output」タブから、道路に対してラベル付けを行うこともできる。

## 5. 運転操作と道路特徴の分析

本稿では、4.節で提案した分析・可視化システムを用いて様々な道路の運転操作と道路特徴の分析を行った。今回は、事故の多くは交差点で発生していることから、対象を交差点に絞った。また、運転操作の少ない交差点では十分な統計が取れないことから、周辺で各運転操作が30以上記録された、350箇所の交差点を対象とした。これらの交差点における交通事故件数を、交通事故データよりカウントした。この結果、350箇所中146箇所では事故が発生していることが分かった。

次節以降では、この350箇所の交差点について運転操作と道路特徴の分析を行った結果を述べる。

### 5.1 運転操作分析

#### 5.1.1 事故の有無による運転操作の分布の比較

まず、交通事故の有無による運転操作の分布を比較した(図6)。赤色が事故があった交差点、青色が事故がなかった交差点での運転操作の分布である。この図から、事故があった交差点では、ブレーキ操作の最大速度、ハンドル操作の最大速度、右左折操作の進入速度、最大速度、終了時速度が高い傾向にあることが分かった。逆に、ハンドル操作の最大角加速度、最大角加速度変化は低い傾向にあることが分かった。ここから、事故の発生した交差点では、速い速度で大きくゆっくりと曲がる傾向にあると考えられる。

#### 5.1.2 事故と道幅の太さによる運転操作の分布の比較

前節では事故の有無によって運転操作の分布に違いが生じることが分かった。ところが、事故の発生した交差点を個々に分



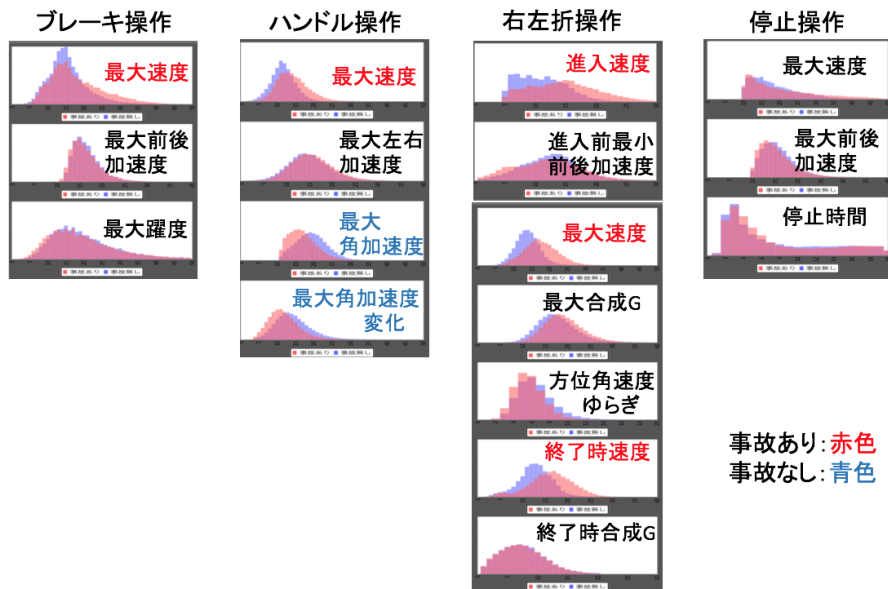


図 6 事故の有無による運転操作の分布の比較

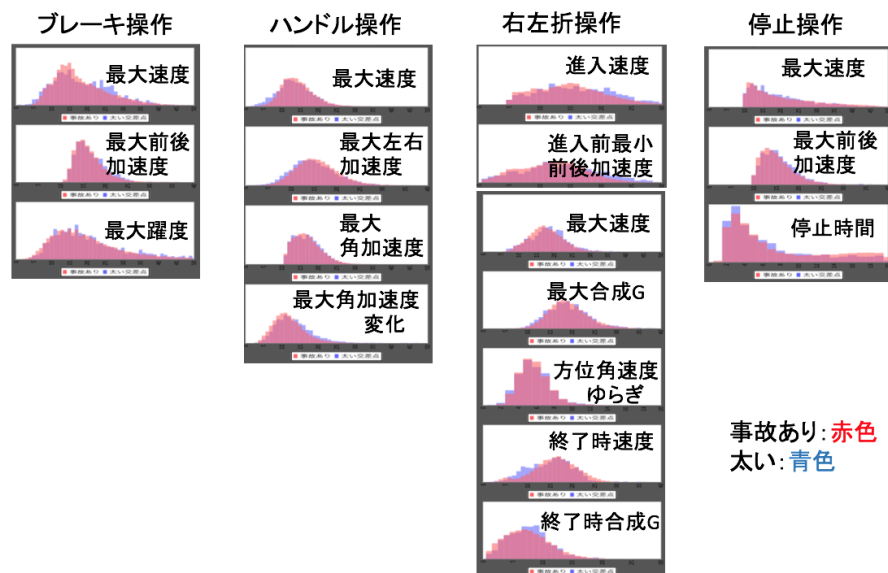


図 7 事故と道幅の太さによる運転操作の分布の比較

析していたところ、道幅が太い交差点で事故が多く起きていることが判明した。そこで、事故が発生した交差点と道幅が太い交差点での運転操作の分布を比較することにした。なお道幅が太い交差点は、ゼンリンの道路地図データより、道幅が13m以上の道路が3本以上接続している交差点と定義した。結果は図7である。各運転操作のどの属性の傾向もほぼ一致していることが分かった。このことから、前節で判明した事故の有無による運転操作の分布の違いは、太い交差点の分布の比較であった可能性が考えられる。

### 5.1.3 道幅の太さによる事故発生数

本節では、道幅の太さによって事故発生数がどう異なるかを調べた結果について述べる。表3に、道幅が太い交差点と細い交差点それぞれの事故発生数をまとめた。なお、道幅が細い交差点は、太い交差点同様に道路地図データを基に、道幅が

13m未満の道路が4本接続している交差点とした。この表から、ほとんどの太い交差点で事故が発生していることが分かった。太い交差点であること自体がリスクであることが判明したため、細い交差点について、次節より詳細な道路特徴の分析を行うことにする。

表 3 道幅の太さによる事故発生数の比較

|       | 0 件 | 1 件 | 2 件 | 3 件 | 4 件以上 |
|-------|-----|-----|-----|-----|-------|
| 太い交差点 | 3   | 6   | 12  | 9   | 55    |
| 細い交差点 | 201 | 52  | 6   | 2   | 2     |

### 5.2 道路特徴分析

本節では、前節で定義した細い交差点について、詳細な道路特徴を分析した結果について述べる。まず対象としたのは、前節で対象としていた350箇所の交差点のうち、263箇所の細い

交差点とした。これらの交差点に対して、分析・可視化システムを用いて道路特徴の有無を調べた。本稿では、システムを用いて道路を探索した結果、事故の要因となりうる道路特徴を選択し、各交差点のストリートビューを見て各特徴の有無を手でラベル付けした。今回選択した道路特徴は、坂道、一時停止、横断歩道、スクールゾーン、歩道、ガードレール、駐車場、見通しの良さの8つである。

表4に、各道路特徴の有無と事故発生率をまとめた。なお、信号と一方通行については、ゼンリンの数値地図データを用いた。カッコ内は各道路特徴を分母としている。また、カイ2乗検定において有意水準1%で有意差があった項目は太字で表記した。この表から、信号、坂道、横断歩道、歩道、ガードレールの5つの道路特徴は、ある場合の方が事故が発生しやすいことが分かった。逆に一時停止は無い場合に事故が発生しやすいことが判明した。

| 表4 各道路特徴の有無と事故発生率 |                    |                   |
|-------------------|--------------------|-------------------|
| 道路特徴              | ありの時の事故率           | なしの時の事故率          |
| 信号                | 45%(21/47)         | 19%(41/216)       |
| 一方通行              | <b>21%(38/179)</b> | 29%(24/84)        |
| 坂道                | <b>58%(50/86)</b>  | 7%(12/177)        |
| 一時停止              | 17%(34/201)        | <b>45%(38/62)</b> |
| 横断歩道              | <b>38%(49/129)</b> | 20%(31/152)       |
| スクールゾーン           | 29%(22/76)         | 21%(40/187)       |
| 歩道                | <b>33%(46/139)</b> | 13%(16/124)       |
| ガードレール            | <b>33%(39/120)</b> | 16%(23/143)       |
| 駐車場               | 19%(21/110)        | 27%(41/153)       |
| 見通し良い             | 29%(39/136)        | 18%(23/127)       |

## 6. 潜在リスク箇所抽出

本節では、分析・可視化システムを用いて運転操作・道路特徴を分析した結果を基に、潜在リスクの存在する交差点の抽出を目指す。ここでの潜在リスクの存在する交差点とは、未だ事故が発生していないが、リスクが高い交差点を指す。まず6.1節では、実際に事故が発生した交差点を判別する分類器の構築、評価を行い、その結果を基に6.2節でリスクが存在する交差点が抽出できたかを考察する。

### 6.1 交通事故判別実験

本節では、既にリスクが存在すると考えられる、事故が発生した交差点を判別することを試みる。特徴量は、運転操作と道路特徴を用いた。運転操作は各操作の各属性ごとの特徴量ヒストグラムをベクトルとして使用した。道路特徴は、前節での10の道路特徴の有無を特徴量とした。判別対象は前節の263箇所の細い交差点とし、このうち事故が発生した62箇所を正例、事故が発生していない201箇所を負例とした。分類器としてはSVM (RBF カーネル)、ライブラリとしてはscikit-learn<sup>(注9)</sup>を用いて10分割交差検定を行い、適合率、再現率、F<sub>1</sub>値を求めた。また、ベースラインとして、全てを正例と判別した場合の結果との比較も行った。

表5 SVMによる事故の有無の判別結果

| 素性         | 適合率  | 再現率  | F <sub>1</sub> 値 |
|------------|------|------|------------------|
| ベースライン     | 0.24 | 1.0  | 0.38             |
| 道路特徴       | 0.40 | 0.74 | 0.52             |
| 運転操作       | 0.48 | 0.63 | 0.54             |
| 運転操作+ 道路特徴 | 0.50 | 0.70 | 0.58             |

分類結果は表5の通りである。運転操作を特徴量としたときの判別性能は、ベースラインよりもおよそ0.17高い精度であった。また、道路特徴のみを使った場合でも0.14高くなり、両方を特徴量とした場合、ベースラインよりも0.2ほど高い精度で判別できた。このことから、運転操作だけでは判別が難しい場合でも、人手によって道路特徴をラベル付けすることで、判別性能の向上が期待できることが分かった。

### 6.2 潜在リスク箇所抽出

表6は、6.1節での判別実験のうち、運転操作と道路特徴を特徴量として用いた場合の詳細な結果である。

表6 運転操作と道路特徴を特徴量とした場合の事故判別結果の詳細

|         | 事故あり | 事故なし |
|---------|------|------|
| 事故ありと判別 | 43   | 43   |
| 事故なしと判別 | 19   | 158  |

表のうち、事故が発生していないが、事故が発生したと判別された43箇所の交差点は、リスクが存在すると考えられる。そこで本節では、これらの交差点のうちいくつかを分析した結果を述べる。分析にはストリートビューと運転操作の分布を用いた(図8, 9, 10)。運転操作の分布は、各交差点での分布(図中の赤色のヒストグラム)と、事故が発生した交差点のうち事故ありと判別された交差点の分布(図中の青色のヒストグラム)を比較した。

1つ目の交差点(図8)は、信号があるが、交差点の角まで建物が建っており見通しが悪い。また、ストリートビューからは交差点の一角にある飲食店に行列ができていることが確認でき、歩道もガードレールもないため運転には注意を要すると考えられる。さらに路側帯にはスクールゾーンが確認でき、通勤通学の時間帯は特に気をつけるべき場所であると思われる。

2つ目の交差点(図9)は、形が少し急で、かつ塀や電柱が交差点の角にあるため非常に見通しが悪く、直前まで道があることすら気づきにくい。その状況を反映してカーブミラーが設置されており、運転に注意が必要であることが分かる。

3つ目の交差点(図10)は、一見すると危険には見えない。信号や歩道、ガードレールも設置されており、見通しも良い。しかし運転操作の分布を見ると、事故の発生した交差点と似た傾向を示していることが分かった。具体的には、右左折操作の進入速度が速く、ハンドル操作の最大角加速度が低い傾向にあった。このことから、一見するとリスクがなさそうな交差点も、運転操作によって抽出できていることが分かった。

(注9) : <http://scikit-learn.org/>



図 8 1 つ目の交差点の Google ストリートビューと運転操作分布



図 9 2 つ目の交差点の Google ストリートビューと運転操作分布



図 10 3 つ目の交差点の Google ストリートビューと運転操作分布



## 7. まとめと今後の課題

本稿では、運送会社のドライブレコーダデータを分析し、潜在リスク箇所を抽出する手法を提案した。まず、運転操作の分布と道路特徴を分析・可視化できるシステムを構築した。このシステムでは、運転操作が集中している箇所を地図上で効率的に探索し、個々の箇所について詳細な分析を行うことができる。またストリートビューを基に、交通事故のリスクとなりえる道路特徴のラベル付けをシステム上で行うこともできる。このシステムを用いて運転操作の分布を分析した結果、事故の発生した交差点と太い交差点では似た運転操作の分布であることが判明した。また道路特徴の分析を行った結果、信号、坂道、横断歩道、歩道、ガードレールの5つは有る場合の方が、一時停止は無い場合の方が事故が発生しやすいことが分かった。次に、事故が発生した交差点を判別する実験を行った。この結果、運転操作と道路特徴を特徴量として用いた場合、ベースラインより  $F_1$  値が0.2向上することを確認した。最後に、事故判別実験において事故が未発生だが発生したと判別された交差点について、潜在リスクの分析を行った。分析の結果、歩道と車道が分かれておらず、スクールゾーンがあるため運転に注意を要する交差点や、塀によって道が見難い交差点など、リスクがあると考えられる道路を見つけることができた。また、一見安全そうに見えるが、運転操作の分布が事故の起きた交差点に類似した道路があることも確認した。今後の課題としては、道路特徴を自動でラベル付けを行うこと、ラベル付けすべき交差点を分析・可視化システムを利用して支援することが挙げられる。また、天候や時間帯、交通量といった情報を考慮することも課題である。

## 謝 辞

本研究の一部は東京大学空間情報科学研究センターとの共同研究に基づき研究用空間データ基盤よりデータ提供を受けて実施している。

## 文 献

- [1] H. W Heinrich, "Industrial accident prevention, A scientific approach," 1941.
- [2] H. Doraiswamy, N. Ferreira, T. Damoulas, J. Freire, and C. T. Silva, "Using topological analysis to support event-guided exploration in urban data," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2634-2643, 2014.
- [3] N. Ferreira, J. Poco, and H. Vo, "Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips," *Comput. Graph.*, vol. 19, no. 12, pp. 2149-2158, 2013.
- [4] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. Van De Wetering, "Visual Traffic Jam Analysis Based on Trajectory Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2159-2168, 2013.
- [5] F. Wang et al., "A visual reasoning approach for data-driven transport assessment on urban road," in *Proc. IEEE Conf. Visual Anal. Sci. Technol.*, pp. 103-112, 2014.
- [6] X. Liang, X. Zheng, W. Lv, T. Zhu, and K. Xu, "The scaling of human mobility by taxis is exponential," *Phys.A,Stat. Mech. Appl.*, vol. 391, no. 5, pp. 2135-2144, 2012.
- [7] H. Liu et al., "Visual analysis of route diversity," in *Proc. IEEE Conf. Visual Anal. Sci. Technol.*, pp. 171-180, 2011.
- [8] J. Pu, S. Liu, Y. Ding, H. Qu, and L. Ni, "T-Watcher: A new visual analytic system for effective traffic surveillance," in *Proc. IEEE 14th Int. Conf. MDM*, pp. 127-136, 2013.
- [9] Y. Liu, C. Kang, S. Gao, Y. Xiao, and Y. Tian, "Understanding intra-urban trip patterns from taxi trajectory data," *J. Geogr. Syst.*, vol. 14, no. 4, pp. 463-483, 2012.
- [10] D. Chu, D. a. Sheets and G. Chen, "Visualizing Hidden Themes of Taxi Movement with Semantic Transformation," *IEEE Pacific Vis. Symposium.*, pp. 137-144, 2014.
- [11] Y. Yue, Y. Zhuang, Q. Li, and Q. Mao, "Mining time-dependent attractive areas and movement patterns from taxi trajectory data," in *Proc. 17th Int. Conf. Geoinf.*, pp. 1-6, 2009.
- [12] Y. Yue, B. Hu, H. D. Wang, and Q. Q. Li, "Identifying shopping center attractiveness using taxi trajectory data," in *Proc. Int. Workshop Trajectory Data Mining Anal.*, pp. 31-36, 2011.
- [13] B. Jiang, J. Yin, and S. Zhao, "Characterizing the human mobility pattern in a large street network," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 80, no. 2, 2009.
- [14] C. A. Quiroga and D. Bullock, "Travel time studies with global positioning and geographic information systems: An integrated methodology," *Transp. Res. C, Emerging Technol.*, vol. 6, no. 1/2, pp. 101-127, 1998.
- [15] G. Andrienko and N. Andrienko, "Spatio-temporal aggregation for visual analysis of movements," *VAST '08*, pp. 51-58, 2008.
- [16] Y. L. Murphey, R. Milton, L. Kiliaris, "Driver's style classification using jerk analysis," *IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, pp. 23-28, 2009.
- [17] D. A. Johnson and M. M. Trivedi, "Driving Style Recognition Using a Smartphone as a Sensor Platform," *IEEE Transactions On Intelligent Transportation Systems*, pp. 1609-1615, 2011.
- [18] W. Wu, W. S. Ng, S. Krishnaswamy, and A. Sinha, "To taxi or not to taxi? - enabling personalised and real-time transportation decisions for mobile users," in *MDM*, pp. 320-323, 2012.
- [19] M. Veloso, S. Phithakkitnukoon, and C. Bento, "Sensing urban mobility with taxi flow," in *Proc of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks*, pp. 41-44, 2011.
- [20] A. Fredrikson, C. North, C. Plaisant, and B. Shneiderman, "Temporal, geographical and categorical aggregations viewed through coordinated displays: A case study with highway incident data," in *Proc. Workshop New Paradigms Inf. Vis. Manipulation*, pp. 26-34, 1999.
- [21] M. L. Pack, K. Wongsuphasawat, M. VanDaniker, and D. Filippova, "Ice- visual analytics for transportation incident datasets," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, pp. 200-205, 2009.
- [22] A. Anwar, T. Nagel, and C. Ratti, "Traffic origins: A simple visualization technique to support traffic incident analysis," *IEEE Pacific Vis. Symposium.*, pp. 316-319, 2014.
- [23] M. Itoh, D. Yokoyama, and M. Toyoda, "Visual Interface for Exploring Caution Spots from Vehicle Recorder big data," in *Proc. Big Data 2015*, pp. 776-784, 2015.