

職業ドライバのテレマティクスデータを用いた急減速情報のクラスタ分析

広野 龍一 指導教員：河中 治樹

1 はじめに

交通事故低減のために様々な対策がなされている。その中に急減速に注目した研究として、西堀らは人身事故の発生数と急減速の発生率の間に関係があることを検証した [1]。また、菊地らは死傷事故率に代替される指標として急減速をもっており、道路の減速ドット施工前後の急減速回数を比較することで、対策事業の効果を検証している [2]。

しかし、急減速は様々な要因や環境下にて発生することが考えられ、状況に合わせて有効な提言を行うことでより交通事故低減につなげることが可能であると考えられる。

そこで本研究では、記録された急減速の情報をもとに急減速を回数だけではなく、発生した環境や要因別に急減速挙動を分類することを目的とする。

2 取得したデータ

本研究ではテレマティクスデータと、ドライブレコーダー画像を取得した。テレマティクスデータは急減速の解析用、ドラレコ映像は結果の解釈に用いる。

テレマティクスは自動車をはじめとした移動体に通信システムを組み合わせ、情報サービスを提供する新しい技術の総称である。走行中の時刻や座標、速度、3軸加速度等を常時記録しており、運転中のドライバの情報が反映されているビックデータとして関心を集めている。特に保険分野での活用が見込まれており、走行距離によってリスクを算出する従来型の保険に対して、急減速を始めとする危険挙動の回数に応じてリスクを算出する保険をテレマティクス保険と呼ぶ。

本研究では車両の点検用に使用される正確なデータが取得できる、OBD2 端子によって取得された速度情報を使用する。急減速の算出法として、常時取得される速度情報から現在の時刻の速度を v_t 、現在の 2 秒前の速度を v_{t-2} とした場合 a が $-0.28G$ を下回った場合に急減速と判定される。

$$a = \frac{(v_t - v_{t-2})}{2} \quad (1)$$

急減速と判定された時刻の前後 5 秒間の速度情報を取得し解析を行った。

また、ドラレコは運転中の映像を常時録画しており、急減速が発生した時刻と対応させることで急減速が発生した際の急減速の要因や、周囲の環境を確認した。



図 1 急減速発生要因 (左:対向車 右:歩行者)

3 提案手法

提案する手法の手順として、まず特徴量の抽出を行う。次に不必要な特徴量を削除するため、相関を算出し確認したうえで主成分分析を行う。抽出した主成分を用いてクラスタリングによって急減速挙動を分けることとする。

取得された速度情報について時系列による変化の情報を取得するために、1次微分である加速度、2次微分である躍度を抽出した。また、加速度が閾値である $-0.28G$ を下回り続けた時間とその傾き、1急減速当たりの躍度のとった範囲も特徴量として抽出し、30次元の特徴量を作成した。ドライバ1名の1か月間の全急減速の結果を図2, 2に示す。

特徴量同士の相関が高い場合、重複する情報が含まれるため不必要な情報を削除することで計算量の軽減を図ることができる。そのため、すべての特徴量の組み合わせについてピアソンの積立相関を算出した。ピアソンの相関係数は式(2)で表される。このとき、 s_x は x の標準偏差を、 s_y は y の標準偏差を、 s_{xy} は x と y の共分散を表す。また、算出した特徴量間の組み合わせを色によって可視化した結果を図2に示す。

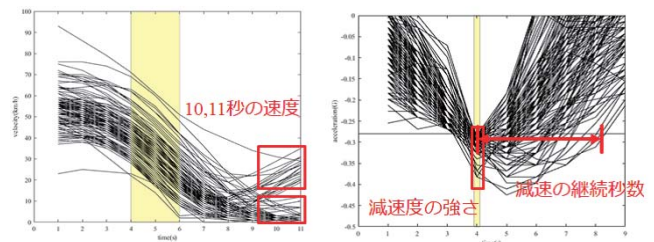


図 2 特徴量の抽出 (速度, 加速度)

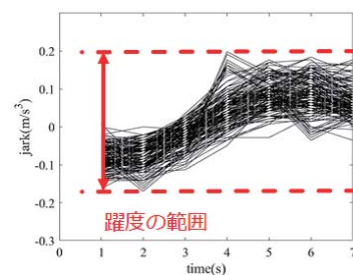


図 3 特徴量の抽出 (躍度)

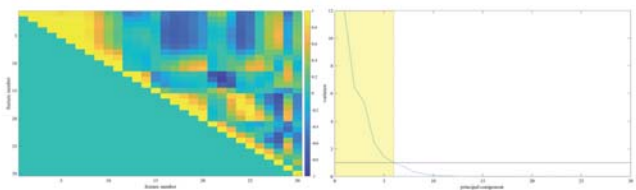


図 4 左:相関 右:主成分ごと寄与率

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{s_{xy}}{s_x s_y} \quad (2)$$

算出した相関をもとに主成分分析を行った。主成分分析は特微量選択の手法の 1 種であり、 n 次元の特微量の共分散が最大となる軸を探す必要がある。

このとき、それぞれの主成分が元のデータをどの程度反映しているかを示す指標として、寄与率が用いられる。寄与率は持つ情報が少ない軸を削除するため、カイザー基準（共分散の固有値が 1 以上）に従い選択した。主成分ごとの寄与率の変化を図 2 に示す。

クラスタリングの手法には階層的なクラスタリングを用いた。非階層的なクラスタリングはデータの分割の結果を評価する関数を作成し、関数を最適にする分割を探索する手法であるが、事前に分割するクラスタ数を決定しなければならずデータの関係が明確になっていない本研究には適していないと考えられる。

一方、階層的なクラスタリングは N 個のデータをそれぞれ 1 個のクラスタとして仮定する。すなわち N 個のクラスタがある場合に、抽出した 2 個のクラスタの距離を算出し、最も距離の近いクラスタを併合する。その後も算出と併合をすべてのクラスタが 1 個のクラスタになるまで繰り返すことで、階層的にクラスタが併合される過程を知ることが可能である。

データ間の距離の算出には多くの手法があるが本研究ではユークリッド距離を使用している。任意の要素 x と y 間の距離は式 (3) によって表される。

また、クラスタ内の重心を求める手法として、鎖効果の発生しにくい群平均法を用いる。データの個数を N 、距離を算出する対象 \mathbf{x}_1 、 \mathbf{x}_2 の間の距離を $d(\mathbf{x}_1, \mathbf{x}_2)$ 、クラスタ C_1 、 C_2 の間の距離を $d(C_1, C_2)$ としたときの距離の導出法を式 (4) に示す。

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (3)$$

$$d(C_1, C_2) = \frac{1}{|C_1||C_2|} \sum_{\mathbf{x}_1 \in C_1} \sum_{\mathbf{x}_2 \in C_2} d(\mathbf{x}_1, \mathbf{x}_2) \quad (4)$$

4 解析の結果

タクシードライバ 6 名（男性 4 名女性 2 名、平均年齢 68.71 ± 1.39 ）の 1 か月のデータを取得した。それぞれのドライバーについて提案手法で記述した処理を行い、クラスタリングを行った。

それぞれのドライバーのクラスタ分類したクラスタごとの要素数を表 1 に示す。一か月間の急減速数の総数に対してほとんどの急減速が含まれる非常に大きなクラスタが 1 つできていることがわかる。6 人のうち 1 人を例として結果を示す。

それぞれの急減速の発生した状況をドラレコで確認したところ、急減速理由の 9 割り近くは信号停止によるものであることがわかった。急減速要因と交通違反の有無を急減速ごとに調べた結果を図 4 に示す。横軸はそれぞれ要因や、交通違反の名前が入る。今回は特に注目する要素のみラベルを表示している。また、縦軸はそれぞれの挙動が発生した回数を表示しており、グラ

表 1 5 クラスタに分割した際のデータ分布

ドライバー	急減速数	クラスタ 1	クラスタ 2	クラスタ 3	クラスタ 4	クラスタ 5
subA	118	1	114	1	1	1
subB	8	1	4	1	1	1
subC	152	1	139	10	1	1
subD	62	2	56	1	2	1
subE	35	1	1	27	5	1
subF	18	12	3	1	1	1

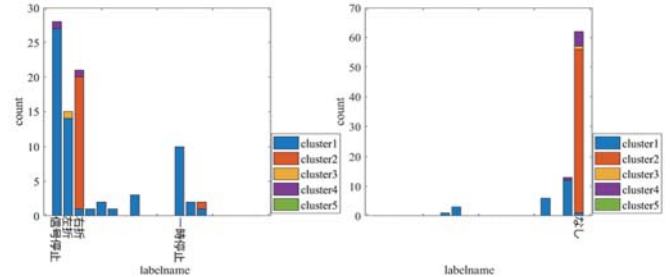


図 5 左:急減速要因 右:交通違反

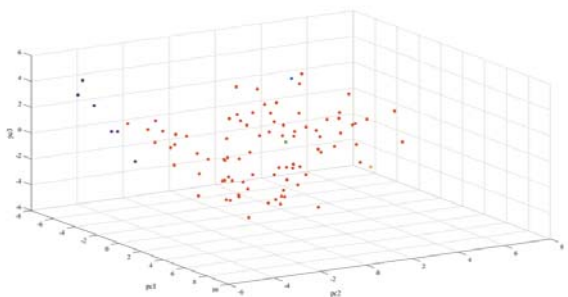


図 6 クラスタ分類された特徴量空間

フにつけられた色はそれぞれの要因ごとに急減速が度のクラスタに属しているかを示している。

急減速が起こった要因について、どのような要因にそれぞれクラスタがどのように分布しているかを調査したところ全急減速理由のうち、クラスタ 2 に属する急減速の要因がおおよそ右折によるものであることが分かった。理由として、右折時には対向車の存在を気にしないとならないため、対向車を待って動き出すまでの時間が他の要因に対して特徴的であることが考えられる。また、右折前の急減速では交差点前で停止することが多いため、交通違反についても違法となる場面が少なることが結果より考察される。

第 1 主成分から第 3 主成分までの特徴量空間について、クラスタがどのように点在しているかを描画した結果を図 4 に示す。急減速の特徴量が空間の中心部に集まっていることが確認できる。現在は速度と速度から算出した特徴量のみを用いており特徴の際が小さいため、速度以外のペダルタイミングやハンドル切り角を用いることでさらにクラスタリングの精度を向上させることが可能であると考えられる。

参考文献

- [1] 西堀泰英ら, 土木学会論文集, vol.42, 2010
- [2] 菊地春海ら, 土木学会論文集, vol.68, No.5, pp.1193-1204, 2012