

ドライバの運転基準の把握に向けた車載センサデータ分析

岩崎賢太¹ 白石陽²

概要: 自動運転車両と手動運転車両が混在する自動運転普及期では、これら2種類の運転基準の違いが問題になる。運転基準とは、車両のブレーキ操作や進入タイミングなど、運転時の判断の指標である。2種類の車両の運転基準が異なることで、自動運転システムによる割り込みや低速走行などの現象が起こる。これにより、手動運転車両は、自動運転システムの運転に対応できずに誤操作による事故や渋滞が起こる可能性が考えられる。したがって、自動運転システムは、周辺を走行する手動運転車両の運転基準を把握し、その運転基準に合わせた運転を行う必要があると考える。そこで本研究では、手動運転車両のドライバの運転基準を把握するために、アクセルやブレーキなどの運転操作の結果である車両挙動の特徴（以下、運転特性）を抽出する。著者らの先行研究では、SAX (Symbolic Aggregate Approximation) に基づいて運転特性を抽出した。その際、SAXによるデータ抽象度が低いほど（つまり表現文字列の種類が多い）、運転特性が有効に抽出できることが示唆された。本研究では、運転特性を抽出するための手法として、データ抽象度に着目し、SAXと非抽象化データによる車両挙動データ分析、比較を行った。本実験では、車両走行中に収集したセンサデータに対して、SAXと非抽象化データによる車両挙動データ分析を行い、ランダムフォレストとSVMによるドライバ分類の評価を行った。その結果、SAXによる分析のドライバ分類のF値が0.925と0.74、非抽象化データによる分析のドライバ分類のF値が0.962と0.875となり、SAXと非抽象化データともにドライバの運転特性を抽出できること、SAXよりも非抽象化データによる分析の方が有効であることが示された。

Analyzing in-Vehicle Sensor Data to Understand Driving Standard

KENTA IWASAKI¹ YOH SHIRAISHI²

1. 背景

自動運転車両は、事故や渋滞の削減に期待されており、自動運転社会の安全と環境負荷の軽減を目的とした自動運転技術の研究開発が活発に行われている[1]。2017年には、無人の自動運転車両の公道走行実験が行われており[2]、2020年の11月には、本田技研工業が、世界初の自動運転レベル3の自動車の認可を受け、特定の周辺環境と範囲内において、前方車両との車間距離を維持しながらの自動走行が可能になった[3]。

しかし、自動運転車両は、短期間で大量生産することが難しいため、長い時間をかけて徐々に普及すると考えられる[3]。そのため、自動運転システムが車両を制御する自動運転車両と人間が車両を制御する手動運転車両の2種類の車両が混在することが想定される。これら2種類の車両が混在している時期を自動運転普及期と呼ぶ。自動運転普及期において、自動運転システムの運転によっては、事故や渋滞が起こる可能性があると考えられる。自動運転システムと人間のドライバとでは、減速や停車の際のブレーキ操作の仕方や交差点に進入するタイミングなど、ドライバが安

全に運転するための判断の指標である運転基準が異なる。

この運転基準の違いは、自動運転システムと人間のドライバだけではなく、人間のドライバそれぞれで異なる。運転基準が異なることで、自動運転システムによる急な割り込みや安全を優先したゆっくりとした走行により、ドライバが自動運転システムの運転に対応できずに誤操作することや交通の流れが悪化することが考えられる。また、運転基準は、時間や季節、周囲の道路交通環境などの周辺環境によって変化し、変化の仕方にも個人差がある。そのため、本研究では、自動運転普及期において、周辺環境を考慮した運転基準を自動運転システムに共有することが重要であると考えられる。

本研究では、ドライバの運転基準は、発進や停車といった運転行動中に行われるアクセルやブレーキなど運転操作の結果である車両挙動の特徴（以下、運転特性）に依存すると考える。著者らの先行研究[5]の実験結果より、SAX (Symbolic Aggregate Approximation) [6]を用いてセンサデータを抽象化する際、データ抽象度を低くするほど、運転特性を効果的に抽出することができた。そのため、本研究では、データ抽象度に着目した車両挙動データ分析と評価を行う。車両挙動データ分析手法として、データ抽象度を任意に設定できるSAXと、非抽象化データによる分析の2種類の分析を行う。本研究では、一つのシナリオの周辺環境を対象に運転特性を抽出しているが、最終的には、天気

*1 公立はこだて未来大学大学院 システム情報科学研究科
Graduate School of Systems Information Science, Future University Hakodate

*2 公立はこだて未来大学 システム情報科学部
School of Systems Information Science, Future University Hakodate

や時間、季節、路面状況や交通量などの考えられる全ての周辺環境に対応した運転特性を抽出し、その運転特性と運転基準を対応づけることで、自動運転システムは、各ドライバの運転基準に合わせた運転が可能になると考える。これにより、自動運転車両と手動運転車両が混在した環境下においても、自動運転システムの運転による渋滞や手動運転車両の急ブレーキや急ハンドルが解消され、安全でかつ快適な運転が可能になると考える。

2. 関連研究

本研究では、データ抽象度に着目した分析を行う。関連研究として、2.1 節では、抽象化データによる走行データの分析の研究について述べる。2.2 節では、非抽象化データによる走行データの分析の研究について述べる。

2.1 抽象化データによる走行データの分析

抽象化データを用いた走行データの分析による研究として、谷口らは、ドライバの車両意図には、二重分節構造を含んでいると仮定し、時系列データからチャンクを抽出する手法を提案している[7]。この手法は、時系列データから隠れた二重分節構造を抽出するには、sticky hierarchical Dirichlet process-hidden Markov model (sticky HDP-HMM) を用いて時系列データのモデリングと分節化を行い、nested Pitman-Yor language model (NPYLM) を用いて時系列データから推定された隠れ状態列から文字列をチャンク化している。しかし、この手法では、データを任意の抽象度でデータを抽象化できないため、抽象度を任意に調整できる手法を検討することとその手法が隠れた二重分節構造が抽出可能であるかについて調査する必要があると考える。横山らは、時系列データを文字列に変換する SAX を用いて車両挙動の種類を推定している[8]。この手法では、まず、スマートフォンを車載して収集したセンサデータに対して、任意に閾値を設定し、時系列データを文字列に変換する SAX を適用することでセンサデータを抽象化している。次に、抽象化したデータに N-gram や BoW (Bag of Words) といった自然言語処理技術を用いて車両挙動の種類を推定している。しかし、N-gram で作成した部分文字列の時間情報が考慮されておらず、抽出した文字列が全体の車両挙動のどの部分に相当するかがわからないため、車両挙動の順序関係が考慮できず、運転特性の一つであるタイミングの特徴が抽出できないと考える。

2.2 非抽象化データによる走行データの分析

非抽象化データを用いた走行データの分析による研究として、赤松ら[9]と岡田ら[10]は、時系列間の類似度を算出する手法である DTW (Dynamic Time Warping) を用いた分析を行っている。赤松らは、発進、停車、右左折に対し

て、OBD2 (On-Board Diagnostic second generation) を用いて加速度データを収集し、DTW を適用することで、ドライバの運転特性を分類している[9]。岡田らは、車載センサから収集したアクセル開度とブレーキ圧に DTW を適用し、アクセルパターンとブレーキパターン頻度の割合を分析することで、各ドライバの運転特性を抽出している[10]。しかし、停車以外の運転行動では、ドライバの運転特性に大きな違いが確認できていない。このことから、車両挙動の変化が小さい運転行動において、同様もしくは類似した運転特性が抽出され、運転基準の把握に影響があることが考えられる。鳥居らは、顔と腕に筋電センサを装着し、視線情報と皮膚情報の筋電位データを k-means 法による分析を行い、追従走行時のギア比特性を識別するのに有効な特徴量を評価している[11]。しかし、筋電センサは、身体に装着する必要があるため、ドライバへの装着負担が大きいと考える。また、走行時のドライバの動きを対象にデータを収集するため、車両ごとのハンドルの重さや加減速の大きさなどの車両特性の違いから、車種によって車両挙動の変化から得られるセンサデータに誤差が生じると考える。小林らは、腕と手に加速度センサを装着し、ハンドル操作の加速度データを収集し特異値分解を適用し、ドライバの類似成分と特徴を表す成分を抽出している[12]。スマートフォンを用いた研究として、林らは、車両に固定したスマートフォンから位置情報と 3 軸加速度などを収集し、SVM (Support Vector Machine) を用いた回帰モデルを構築することで、コーナリングにおけるドライバの運転特性を抽出している[13]。

2.3 まとめ

本研究では、データ抽象度に着目した分析を行うため、2.1 節では、抽象化データを用いた走行データの分析による関連研究、2.2 節では、非抽象化データを用いた走行データの分析による関連研究について述べた。

抽象化データによる走行データの分析では、谷口ら[7]は、時系列データからのモデリングと分節化を行い、時系列データから推定される隠れ状態列から文字列をチャンク化している。横山ら[8]は、時系列データを文字列に変換する SAX を適用することで、データを抽象化している。本研究では、データ抽象度に着目した分析を行うため、任意の抽象度でデータを抽象化する必要がある。そのため、任意に閾値を定めることが可能な SAX の方が適していると考え、抽象化データによる分析では、SAX を用いる。

非抽象化データによる走行データの分析では、赤松ら[9]と岡田ら[10]は、DTW を用いて時系列データを分析することで、ドライバの運転特性やアクセル、ブレーキのパターンを抽出している。鳥居ら[11]は、筋電センサを用いて、追従走行時のギア比特性の有効な特徴量を抽出している。小林ら[12]は腕と手に取り付けた加速度センサに特異値分

解を適用している。林ら[13]は、車両に固定したスマートフォンからセンサデータを収集し、回帰モデルを構築している。

本研究での抽象化データは、非抽象化データに対して SAX を適用することで作成する。ここで、重要となるのは、ドライバの運転特性を抽出するための有効なデータ収集手法である。文献[11][12]のように、ドライバの身体にデバイスを装着する方法では、ドライバへの負担や車両特性によるセンサデータの誤差により、運転特性が有効に抽出できないと考える。文献[9][10][13]では、CAN (Controller Area Network) や OBD2, スマートフォンを用いて車両情報もしくは、車両挙動の変化に関するセンサデータを収集している。これらは、身体にデバイスを装着する方法と比較して、ドライバの意思決定から実際に運転操作するまでの時間間隔が大きい、ドライバへの負担が少なく、車両特性による影響も小さい。そして、この時間間隔は、ドライバの運転特性に大きな影響はないと考える。CAN や OBD2 は、車両情報や車両挙動の変化に関するデータが収集できるが、設置箇所が変更できない。一方で、スマートフォンは、アクセル開度や車両速度、エンジンの回転数などの車両情報は取得できないが、車両挙動に関するデータの収集と任意の場所に設置できる。著者らの先行研究[5]や文献[8][13]では、スマートフォンに内蔵されているセンサデータを分析し、目的の対象を分類、推定している。さらに、後部座席や助手席など複数箇所に設置可能なため、より詳細に車両挙動のデータを収集できると考える。そのため、本研究では、スマートフォンを用いたデータ収集を行い、抽象度別のデータ分析を行う。

3. 提案手法

3.1 研究目的

本研究の目的は、ドライバの運転基準を把握するために、車両挙動からドライバの運転特性を抽出することである。本稿では、車両挙動として「停車」、「右折」、「左折」を取り上げ、これらの車両挙動の分析結果について述べる。1章で述べたように自動運転システムと人間のドライバとでは、運転基準が異なるため、自動運転システムの運転によっては、人間のドライバは「危険」を感じることがあり、事故や渋滞につながる事が考えられる。そのため、自動運転システムが周辺ドライバの運転基準に合わせて運転することが必要である。本研究では、ドライバの運転基準は、運転特性に依存すると考えるため、車両挙動に関するデータを収集、分析することで、運転特性を抽出する。著者らの先行研究では、SAX を適用し、データを抽象化してから運転特性を抽出した。抽出した運転特性を評価した際、データの抽象度を低くするにつれて、効果的に運転特性を抽出できることが示唆された。そこで本研究では、データ抽

象度に着目し、車両挙動に関するデータに対して、任意に抽象度でデータを抽象化可能な SAX と非抽象化データによる分析を行うことで、運転特性を抽出する。

3.2 研究課題とアプローチ

3.1 節で述べた研究目的を達成するには、以下の 3 つの課題を解決する必要がある。

- a) 運転特性の抽出に有効なデータ収集方法の検討
- b) SAX を用いた分析による運転特性の抽出
- c) 非抽象化データを用いた分析による運転特性の抽出

課題 a のアプローチとして、車載したスマートフォンに内蔵されているセンサから車両挙動に関するデータを収集する。2.2 節、2.3 節で述べたように、装着型センサを用いた場合は、データ収集時のドライバへの装着負担が大きい。さらに、車両の形状、ハンドルの重さや加減速の大きさといった車両特性が異なる。そのため、収集したデータと実際の車両挙動の変化が一致するとは限らない。したがって、分析結果に誤差が生じる原因になると考える。車載センサを用いた場合は、ドライバにデバイスを装着する必要がないため、ドライバへの装着負担がない。また、アクセル開度やブレーキ圧、エンジンの回転数などの車両情報もしくは、運転操作の結果である車両挙動を収集するため、装着型デバイスよりも車両特性による影響が小さいと考える。車載センサを用いた方法としては、CAN とスマートフォンを用いた方法が挙げられるが、最終的な運転操作結果の車両挙動の収集を目的としているため、車両特性の影響が小さい、ダッシュボードに設置したスマートフォンを用いて車両挙動データを収集する。

課題 b では、「停車」、「右折」、「左折」のように、運転行動毎のセンサデータに SAX を適用し、ランレングス符号化、文字変化割合などの文字列処理を用いて運転特性を抽出する。ランレングス符号化を行うことで、文字のみで構成される文字成分と数字のみで構成される数値成分に分類することで、車両挙動の変化と車両挙動の変化に要する時間に分割する。さらに、時系列情報を付与するために、数値成分を基に算出した累積和。これら 3 つを分析することで、ドライバの運転特性を有効に抽出できると考える。また、文字変化割合を求めることで、文字の変化のパターンとその文字の変化量から、運転特性を有効に抽出できると考える。SAX を用いた車両挙動データ分析手法の詳細は、3.4 節で説明する。

課題 c では、DTW 距離と最大値、最小値、標準偏差、中央値などの統計量を算出することで、運転特性を抽出する。まず、DTW 距離は、一つの基準データと各試行分の車両挙動データとの DTW 距離を算出し、その平均を分析する。これにより、運転操作による車両挙動の変化のばらつきが

把握できると考える。続いて、統計量については、ドライバの運転操作の特徴を分析する。最大値と最小値では、ドライバのペダルの踏み方やハンドルの回し方の特徴が抽出できると考える。このように、運転の安定度と運転操作の特徴の2つを分析するために、DTW 距離の算出と統計量を用いた分析が有効であると考えられる。

3.3 車両挙動データ収集

本研究では、車載ホルダーを用いてスマートフォンを車両のダッシュボードの上に設置し、車両挙動データを収集する。まず、車両の動きや傾き、振動などが計測できるセンサデータとして、加速度センサ、角速度センサ、地磁気センサからセンサデータを取得する。収集するセンサデータは、サンプリングレート 50Hz の 3 軸加速度センサ、3 軸角速度センサ、3 軸地磁気センサとサンプリングレート 10Hz の GPS による位置情報である。データ収集時のスマートフォンは、各 3 軸センサの X 軸が車両の左右の方向、Y 軸が鉛直方向、Z 軸が進行方向となるようにダッシュボード上に設置した。また、データ収集にあたって、ドライバの運転操作のタイムスタンプを取得するために、ダッシュボードに設置したスマートフォンとは別に、他 2 台のスマートフォンを用いて、ハンドル操作とペダル操作を撮影した。センサデータ収集環境とセンサ軸の向きについて図 1 に示す。

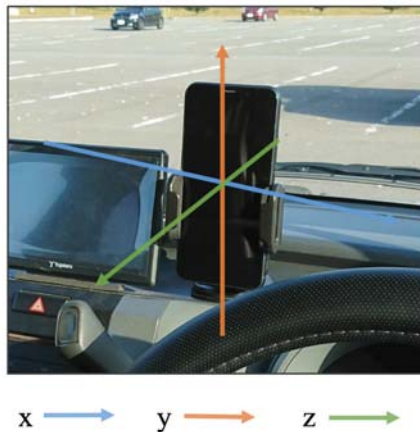


図 1 スマートフォンの設置環境と各センサ軸の向き

3.4 SAX による車両挙動データ分析

本節では、収集したセンサデータから運転特性を抽出する手法について述べる。3.4.1 項で SAX によってセンサデータを文字列に変換する流れについて述べる。3.4.2 項では、ランレングス符号化について述べる。3.4.3 項では、基準文字による部分文字列の作成について述べる。3.4.4 項では、文字変化割合について述べる。

3.4.1 SAX によるセンサデータの文字列変換

SAX は時系列データを文字列に変換する手法で、時系列データのノイズ除去やデータ容量が削減に期待されている[14]。本研究では、SAX を適用する前処理として、センサデータの標準化を行う。SAX によって時系列データを文字列に変換する手順を以下に示し、そのイメージを図 2 に示す。

- (1) 時系列データが正規分布に従っていると仮定し時系列データを標準化（平均が 0、分散が 1 になるように変換）する。
- (2) 時系列データを時間軸に対して等間隔に分割する。
- (3) 分割された時間内のデータの平均値を算出する。
- (4) 正規分布にしたがって分割した区間に文字を割り当てる。
- (5) 手順(3)より算出した値を手順(4)に基づいて文字に変換する。

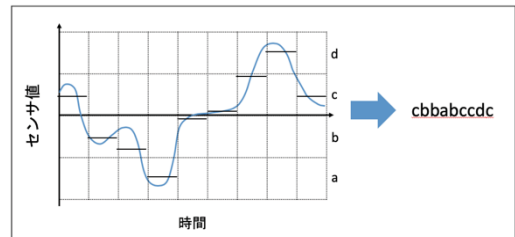


図 2 SAX のイメージ

本研究では、データ抽象度に着目した分析を行う。本提案手法では、正規分布にしたがって閾値を定めた。SAX は、この閾値を一定間隔に設定することや、割り当てる文字を増やすことで、閾値の間隔を小さくすることができる。これにより、データの抽象度を設定する。本研究では、収集したセンサデータを 4, 8, 16 種類の文字で表現する。このセンサデータを表現する文字の種類を表現文字種類 W とする。

3.4.2 ランレングス符号化による車両挙動の表現

時系列データに SAX を適用して変換した文字列をランレングス符号化する。3.4.1 項でセンサデータを文字列に変換する手順について述べた。SAX によって変換した文字列は、センサデータ値の変化情報と時間情報の二つを含んでいる。そこで、ランレングス符号化を行い、文字成分と数値成分に分けることで、車両挙動データ分析を容易にする。ここで、ランレングス符号化は、元の文字列の情報を完全に保持したまま、文字列を圧縮できるため、ランレングス符号化された文字列（以下、ランレングス符号化文字列）から元の文字列への復元が可能である。よって、SAX によって変換した文字列に、ランレングス符号化を適用しても、運転特性の抽出の影響は少ないと考える。ランレングス符号化文字列は、文字と数値が交互に並んで構成されており、

各数値部分に対してそれまでの数値成分の累積和を計算し記録することで、先頭文字からの経過時間を表現する。

3.4.3 基準文字による部分文字列の抽出

累積和を用いた運転特性の特徴を抽出するために、基準文字を設定し、ランレングス符号化文字列から抽出した文字成分を基準文字に従って分割する。分析対象の運転行動としては発進、停車、右左折などがあるが、各運転行動によって、運転操作に対する車両挙動データの変化の仕方が異なるため、運転操作に対応する部分文字列の作成が必要である。そのため、各運転行動と各センサデータに応じて基準文字を設定する必要がある。ここでは、停車の進行方向の加速度を例に部分文字列を作成手順について説明する。まずは、基準文字を設定する。加速度センサや角速度センサは、静止している状態の値が0に収束する。さらに、本研究の表現文字種類 W は、4, 8, 16 と、全て偶数であることから、0 が閾値となる。そのため、基準文字として、0 に最も近い正の領域の文字を設定する（図2で例えると、表現文字種類 $W=4$ のため、基準文字はCとなる）。次に、設定した基準文字が現れた箇所から、次に基準文字が現れる箇所までの文字列を抽出することで、部分文字列を作成する。最後に、作成した部分文字列に対して、各成分の対応付けを行う。対応付けを行うことで、ランレングス符号化文字列の情報を部分文字列に反映できる。これにより、それぞれの運転操作に対応した部分文字列を作成し、運転特性であるタイミングの特徴が抽出できると考える。この一連の処理を図3に示す。

文字成分	C	B	A	B	C	C	D	C	C	B	C
数値成分	1	2	1	1	2	2	1	1	1	1	1
累積和	1	3	4	5	7	7	8	9	9	10	11

図3 基準文字による部分文字列の抽出

3.4.4 文字変化割合に車両挙動データ分析

抽出した部分文字列の文字成分の文字変化割合を分析することで、ドライバの運転操作の特徴が抽出できると考える。各部分文字列の文字成分と数値成分を用いて、部分的な運転操作による車両挙動の文字変化割合を分析することで、ドライバの運転操作の特徴を抽出する。ここで、文字変化割合を算出する方法について述べる。まず、SAX適用後の文字列長を L ，“aa”や“ab”といった文字列中の2文字を i, j とし、その文字の出現数を S_{ij} としたとき、文字変化割合 P_{ij} を以下に記す。

$$P_{ij} = \frac{S_{ij}}{L-1} \quad (1)$$

ここで、文字変化割合は、ランレングス符号化で抽出した文字成分のように同じ文字が連続しない文字列を対象にしても文字変化割合を算出できるが、数値成分の情報が欠落することが考えられる。そのため、“aa”といった同じ文字が連続する文字も対象にすることで、数値成分の情報も文字変化割合に与えることができる。各車両挙動データに上記の式を算出することで、ドライバのブレーキの強さやハンドル操作などの特徴を抽出する。

さらに、各部分文字列には、累積和の情報を付与しているため、運転操作のタイミングと運転操作の時系列が把握できる。停車を例にすると、ブレーキを始めた位置でやや強めにブレーキを踏み、徐々に減速したドライバと停車位置の直前にブレーキを強く踏んで減速・停車したドライバの違いが抽出できると考える。

3.5 非抽象化データによる車両挙動データ分析

本節では、データを抽象化せず、非抽象化データから車両挙動データ分析を行う手法について述べる。非抽象化データによる車両挙動データ分析では、運転の安定度と運転操作の2つの観点から分析を行う。まず、運転の安定度を分析するために、DTW距離を算出する。DTW距離を算出する方法として、収集した車両挙動データのうちの一つを基準データとし、それ以外の各車両挙動データとのDTW距離を算出する。さらに、算出したDTW距離の平均を求めることで、ドライバの運転の安定度を把握する。続いて、ドライバの運転操作を分析するために、統計量を用いる、用いる統計量として、最大値、最小値、平均、中央値、標準偏差、歪度、尖度の6つである。最大値と最小値では、ドライバのペダルの踏み方やハンドルの回し方、歪度や尖度からは、曲がり始めや終わり時のハンドルの回し方や曲がる際の曲率の大きさの特徴が抽出できると考える。このように、運転の安定度と運転操作の特徴の2つを観点から、各特徴量を算出して、ドライバの運転特性を抽出する。

4. 実験

本実験では、SAXと非抽象化データ、2種類の車両挙動データ分析によって、抽出したドライバの運転特性の有効性について評価、比較を行った。

4.1 データ収集における周辺環境

ドライバの運転は、時間、季節、天気、交通量、路面状況、道路形状などの周辺環境に影響される。そのため、周辺環境によるドライバの運転特性の変化が考えられる。したがって、本研究では、ドライバの運転特性を有効に抽出するために、データ収集における周辺環境の構成要素を明確に設定し、各周辺環境の分析が必要であると考え。ここで、データ収集における周辺環境を、車両挙動、道路環

境、道路以外の環境の3つに分け、それぞれの構成要素の検討を行った。データ収集における、各周辺環境の要素の検討結果を表1に記す。

表1 各周辺環境の構成要素

車両挙動		
直進	発進	停車
加速	減速	右左折
カーブ(右)	カーブ(左)	車線変更
道路環境		
交通量	車間距離	周辺車両数
道路形状	車線数	車線幅
勾配	曲率	路面の材質
路面の凹凸		
道路以外の環境		
天気	時間	季節
視界		

表1の周辺環境全てを想定した場合、車両挙動データを分析する際に、考慮する項目が多い。そのため、車両挙動データの分析手法の土台を築くための周辺環境を設定した。

まず、車両挙動として停車、右折、左折を設定した。これは、文献[9]より、停車時のブレーキ操作は、他の車両挙動と比べて個人差が出やすいからである。右折と左折は、基本的な車両挙動であり、曲率やペダル操作に個人差が現れやすいと考える。よって、車両挙動として停車、右折、左折に着目した。次に、道路環境では、交通量がなく、道路形状が直線で勾配がない道路環境とした。これは、ごく普通の道路であり、基本となるドライバの運転特性が抽出できると考える。最後に、道路以外の環境に対しては、天気は晴れ、時間は昼、季節は、春から秋(積雪量0)、視界は良好と設定した。これは、大半のドライバが運転しやすい環境であると考えたためである。道路環境の要素と同様に、基本となるドライバの運転特性が抽出できると考える。以上の設定した要素を仮の周辺環境とし、実験、評価を行った。

4.2 停車と右左折を含む走行実験

本実験では、停車、右左折の車両挙動データを収集するため、普段車を運転している4人の被験者に対して、走行実験を行った。

実験手続きとして、実験者は、まず、被験者に走行コース、停車位置、車両が停車した際の運転の説明を行った。次に、実験者はセンサデータの計測と同時にスマートフォンの録画を開始し、被験者の手元と足元が映るように、スマートフォンを設置した。その後、被験者に指定したコースを走行するように教示した。実験者は被験者がコースを

走行している間、ブレーキ操作のタイムスタンプを収集するために、被験者の足の動きを確認しながら、タイムスタンプを記録した。続いて、被験者は、実験者による実験内容の説明を受けた後、車両の感覚に慣れるために、コースを走行、停車する練習を行った。本番では、実験者による合図を受けた後、図4の橙色の矢印、青色の矢印の通りに走行した。



出典：国土地理院撮影の空中写真（2011年撮影）を加工して作成

図4 走行ルート

走行実験中、右折と左折停車について、運転操作や車両速度が一定の走行になるように条件を課した。右折と左折の走行条件を以下のように定めた。

- (1) 時速30kmの走行を維持する。
- (2) 右折、左折時、ブレーキ操作を行い、減速してからハンドル操作を行う。
- (3) 右折、左折後、時速40kmになるまで加速する。

続いて、停車の走行条件を以下のように定めた。

- (1) 時速40kmの走行を維持する。
- (2) 図4の赤星地点で停車するように、任意の地点からブレーキ操作を行う。
- (3) 車両を1秒以上完全停車する。

以上の走行条件を被験者に課し、図4の橙色矢印と青色矢印の1往復を1試行とし、各被験者に対して20試行分に、DTW距離を算出するための基準となるデータの1試行分を追加し、合計21試行分を行った。本実験は、基本となる運転特性を抽出するために、4.1節で述べたとおりの周辺環境下で走行実験を行った。そのため、走行実験中に別車両や軽車両が走行していた場合、その試行分のデータを破棄し、再度1試行分を行った。ここで、ドライバの運転には、個人差があることから、実験者は、被験者の運転を停車、右左折、運転のばらつきについて評価した。走行実験における、各被験者の運転傾向を表2にまとめた。

表2 各被験者の運転傾向

被験者	ブレーキ属性	右左折	バラつき
P ₁	やや緩やか	スムーズ	なし
P ₂	強め	テキパキ	ややあり
P ₃	とても緩やか	慎重	ややあり
P ₄	とても緩やか	慎重	右左折に バラつきあり

4.3 SAXの分析による運転特性抽出の評価

本実験は、3.4節のSAXによる車両挙動データ分析手法をもとに、収集した車両挙動データの分析を行い、ランダムフォレストとSVMを用いて、ドライバの分類精度による運転特性の抽出の評価を行った。本実験は、著者らの先行研究[5]で行った実験内容を一部変更して行っている。よって、本項では、本実験と先行研究の実験内容と実験結果の比較と考察を行う。まず、本実験内容と先行研究の実験内容について表3にまとめる。

表3 実験内容の比較表

項目	本実験	先行研究
被験者数	4名	4名
運転行動	停車, 右左折	停車
試行回数	20	10
データ数	240	40
特徴量次元数	780	260

表3から、本実験と先行研究の実験の被験者数は、同じではある。しかし、本実験の被験者4名のうち2名は、先行研究の実験の被験者とは異なる。また、本実験では、先行研究の課題となっていたデータ不足と右左折を含めたデータ収集について取り組み、試行回数を倍に増やし、データ数と特徴量次元数が、先行研究の6倍となっている。さらに、全ての特徴量が平等に学習されるために、特徴量の値を標準化している。この他の走行環境、走行ルート、路面状況、走行指示などは、本実験と先行研究と変わらない。

次に、本実験と先行研究のドライバ分類した結果のF値のグラフを図5に記す。

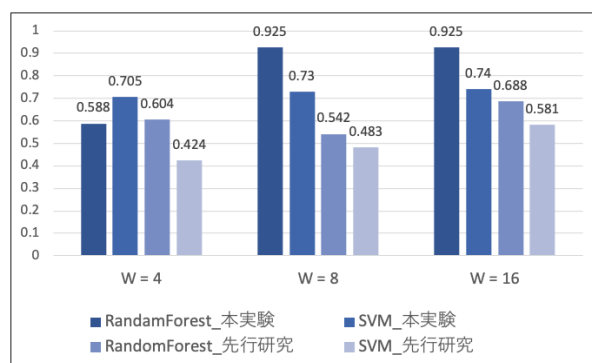


図5 本実験と先行研究のF値

図5より、表現文字種類 $W=4$ のときのランダムフォレストによる分類以外では、本実験のドライバ分類精度が先行研究よりも2割から4割程度高い結果となった。これは、右左折のデータが加わり、データ数が増えたことで、ドライバ分類精度が向上したと考えられる。しかし、表現文字種類 $W=4$ のとき、本実験と先行研究の実験ともにドライバの分類精度が低い。両者のランダムフォレストによる分類では、6割以下、先行研究のSVMによる分類に関しては、4割程度の分類精度である。このことから、データ抽象度が大きい $W=4$ では、運転特性の抽出に適していないと考える。しかし、本実験では、平坦で乾燥した道路で走行実験を行った。SAXを用いた理由の一つとして、ノイズ除去がある。本実験の走行環境下では、SAXのノイズ除去の効果が現れるほどのノイズが生じなかったことが考えられる。そのため、表現文字種類 W が大きくなっても、本実験と先行研究の実験ともに、ドライバの分類精度が十分に向上していることから、ノイズによる影響は小さいと考えられる。ただし、道路が凹凸、砂利、割れなどによって荒れていた場合、これらによる振動がノイズとして、センサデータに現れることが考えられる。このときに、SAXがノイズ除去手法として有効であるか、どの程度の抽象度にするかで、運転特性を有効に抽出できるかなどの調査が必要であると考える。

4.4 非抽象化データの分析による運転特性抽出の評価

本節では、3.5節の非抽象化データによる車両挙動データ分析手法をもとに、4.2節で用いた車両挙動データに対して、分析を行い、ランダムフォレストとSVMを用いて、ドライバの分類精度による運転特性の抽出の評価を行った。ランダムフォレストとSVMによるドライバ分類結果のF値を図6に示す。

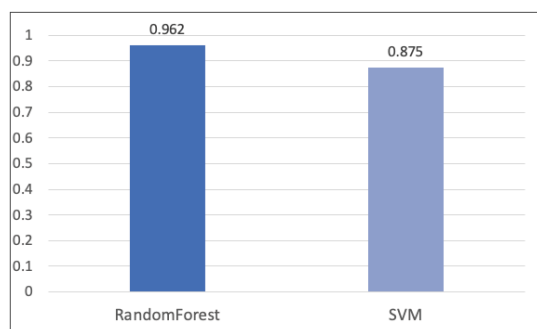


図 6 非抽象化データによる分析のドライバ分類結果

図 6 より、ランダムフォレストでは、ドライバ分類精度が 0.962, SVM では、0.875 となった。SAX によるドライバ分類精度は、ランダムフォレストが 0.925, SVM が 0.74 である。このことから、SAX より非抽象化データによる車両挙動データ分析の方が、ドライバの運転特性をより有効に抽出できたと考える。

4.5 混同行列による SAX と非抽象化データによる分析の比較

本節では、SAX による分析と非抽象化データによる分析を比較するために、各分析の SVM を用いた際のドライバの分類結果の混同行列を算出した。その混同行列を表 4、表 5 に記す。

表 4 SAX による分析の混同行列

		予測				
		P ₁	P ₂	P ₃	P ₄	再現率
正解	P ₁	17	1	1	1	0.850
	P ₂	4	16	0	0	0.800
	P ₃	1	1	10	8	0.500
	P ₄	0	1	5	14	0.700
	適合率	0.773	0.842	0.625	0.609	F 値 0.740

表 5 非抽象化データによる分析の混同行列

		予測				
		P ₁	P ₂	P ₃	P ₄	再現率
正解	P ₁	17	3	0	0	0.850
	P ₂	0	20	0	0	1.000
	P ₃	0	1	16	3	0.800
	P ₄	0	0	2	18	0.900
	適合率	1.000	0.833	0.889	0.857	F 値 0.875

表 4 と表 5 より、被験者 P₁ と被験者 P₂, 被験者 P₃ と被験者 P₄ の 2 つの属性に分類されていることがわかる。この

2 つの属性に対して、被験者 P₁ と被験者 P₂ の属性を属性 1 と被験者 P₃ と被験者 P₄ の属性を属性 2 とする。SAX と非抽象化データともに、属性 1 と属性 2 の 2 つの属性に、ドライバが分類されている。SAX では、属性 1 のデータが属性 2 に分類されている。もしくは、属性 2 のデータが属性 1 に分類されているように、異なる属性に分類された回数が、非抽象化データによる分析と比較して多い。その一方で、非抽象化データでは、同じ属性に属している被験者 P₁ と被験者 P₂, 被験者 P₃ と被験者 P₄ が正しく分類できている。このことから、SAX より非抽象化データの方が、ドライバの運転特性をより有効に抽出できていると言える。ここで、SAX は、閾値を設定し、文字を割り当てるため、データが抽象化される。4.2 節でも述べたように、平坦で乾燥した道路では、SAX のノイズ除去の効果が現れるほど、センサデータには、ノイズが生じなかったことが考えられる。さらに、SAX によるデータの抽象化が、ドライバの個人差、特徴になる部分を丸めていると推測する。そのため、ドライバの分類精度が非抽象化データと比較して劣ると考える。したがって、ドライバの個人差の特徴を抽出する場合は、非抽象化データによる分析を行い、ドライバの個人差を反映させるべきではない対象（車両挙動、路面状況など）の場合は、SAX による分析が適していると考えられる。

5. まとめ

本研究の目的は、ドライバの運転基準を把握するために、収集した車両挙動データからドライバの運転特性を抽出することである。本論文では、著者らの先行研究の実験結果より、データ抽象度に着目した車両挙動データ分析について述べた。本研究では、任意の抽象度でデータを抽象化できる SAX と非抽象化データによる分析の 2 つの分析を行い、ランダムフォレストと SVM を用いたドライバの分類精度による比較を行った。

本実験では、「停車」、「右折」、「左折」の 3 つの運転行動を対象に、車両挙動データを収集し、SAX と非抽象化データによる分析を行い、運転特性を抽出し、ドライバ分類による評価を行った。SAX による分析のドライバ分類の F 値は、ランダムフォレストが 0.925 ($W=16$), SVM ($W=16$) が 0.74 となった。また、表現文字種類 $W=8$ 以上のとき、先行研究の実験結果と比較して、2 割から 4 割程度の精度向上を果たしており、本実験の SAX を用いた車両挙動データ分析が、ドライバの運転特性が抽出できていることが示された。また、SAX を用いたドライバ分類では、先行研究のように、表現文字種類を大きくするにつれて、ドライバ分類精度が向上していることを改めて確認した。非抽象化データによる分析のドライバ分類の F 値は、ランダムフォレストが 0.962, SVM が 0.875 となり、SAX のドライバ分類結果と比べて、1 割程度の分類精度が向上した。さらに、

SAX による分析の混同行列と非抽象化データによる分析の混同行列を比較したところ、両手法ともに、被験者 P_1 と被験者 P_2 、被験者 P_3 と被験者 P_4 の 2 つの属性にドライバ分類されていた。その中でも、SAX は、テストデータと異なる属性に分類していた。一方の非抽象化データでは、テストデータを正しい属性に分類していた。このことから、非抽象化データの方が、ドライバの運転特性をより有効に抽出できていることが示された。

本研究では、平坦で乾燥した道路で走行実験を行い、抽象度に着目した車両挙動データ分析によってドライバの運転特性を抽出した。SAX は、エンジンや路面などの振動によるノイズを除去することを一つの理由として用いていたが、本実験の走行環境下では、SAX のノイズ除去の効果が現れるほどのノイズが生じなかったことが考えられる。また、表現文字種類 $W=4$ のように、データ抽象度を大きく設定した場合、ドライバの個人差を丸めてしまい、運転特性の抽出に悪影響をもたらしていたと考える。そのため、車両挙動や路面状況の推定のような、ドライバの個人差を反映させるべきではない対象では、SAX による分析が適していると考えられる。その反対に、ドライバの個人差のような、細かい変化の特徴を抽出する場合は、非抽象化データによる分析が適していると考えられる。

今後の課題として、ドライバの運転基準と運転特性の対応づけ、運転属性に基づいたクラスタリング、周辺環境の設定の 3 つが挙げられる。

まず、本研究の狙いの一つに、ドライバの運転基準を把握することが挙げられる。そのためには、ドライバの運転基準を測定し、ドライバの運転基準と運転特性を対応づける必要がある。ドライバの運転基準を測定する方法として、VR シミュレータによる運転基準の測定を検討している。実環境で実験を行う場合、同一環境、周辺環境を再現することが困難である。そこで、シミュレータを用いることで、時間、天気、路面状況、交通量などの走行環境を容易に再現可能である。さらに、VR を取り入れることによって、シミュレータによる運転をより現実に近いとできると考える。

次に、運転属性に基づいたクラスタリングについて検討する必要がある。前述したドライバ分類は、クラスタリングによる分類ではない。よって、運転特性を抽出したとしても、その運転特性がどの運転属性に該当するかが識別できないため、運転基準との対応づけができない。そのため、運転属性に基づいたクラスタリングを行う必要がある。

最後に、周辺環境の設定では、ドライバの運転に影響を与える潜在的な要素を検討する必要がある。本研究では、日中で、積雪量、交通量がなく、直線で勾配がない乾燥した道路という 1 種類の周辺環境下で走行実験を行った。ドライバの運転特性は周辺環境に大きく影響を受ける。そのため、周辺環境の設定やドライバと自動運転システムの関

係などを明確にするため、ドライバの運転に影響を与えると考えられる周辺環境を想定する必要があると考える。

謝辞 本研究の一部は JSPS 科研費 JP20K11793 の助成を受けたものである。

参考文献

- [1] 青木啓二：自動運転車の開発動向と技術課題：2020 年の自動化実現を目指して、情報管理, Vol.60, No.4, pp.229-239 (2017).
- [2] Timothy B. Lee : Waymo makes history testing on public roads with no one at the wheel, <<https://arstechnica.com/cars/2017/11/fully-driverless-cars-are-here/>> (参照 2021 年 4 月)
- [3] 国土交通省：報道発表資料 - 世界初！自動運転車（レベル 3）の型式指定<https://www.mlit.go.jp/report/press/jidosha08_hh_003888.html> (2021 年 2 月 2 日参照)
- [4] 辻野照久, 坪谷剛：自動運転自動車の研究開発動向と実現への課題, 科学技術動向, Vol.133, No.1-2, pp.9-16 (2013).
- [5] 岩崎賢太, 白石陽：SAX を用いたブレーキ操作時の運転特性の抽出, マルチメディア, 分散, 協調とモバイルシンポジウム論文集, Vol.2020, pp.1111-1119 (2020).
- [6] Lin, J., Keogh, E., Lonardi S., et al.: A Symbolic Representation of Time Series, with Implications for Streaming Algorithms, SIGMOD Workshop, pp.2-11 (2003).
- [7] 谷口 忠大, 山下 元気, 長坂 翔吾, 坂東 啓司, 竹中 一仁, 人見 謙太郎：運転時系列のベイズ二重分解によるチャート抽出, 計測自動制御学会論文集, Vol.49, No.11, pp.1047-1056 (2013).
- [8] 横山 達也, 白石 陽：交通状況把握のための運転行動の時系列性を考慮した車両挙動分析手法, 情報処理学会論文誌, Vol.60, No.1, pp.87-100 (2019).
- [9] 赤松 伸樹, 河辺 徹, 合原 一究：縦加速度と横加速度の時系列データに基づく乗り心地指標を用いた運転特性解析, 第 61 回自動制御連合講演会, pp.1292-1297 (2018).
- [10] 岡田 将吾, 人見 謙太郎, ナイワラ P. チャンドラシリ, 呂 有為, 新田 克己：車載センサログの時系列データマイニングに基づく運転挙動の分析, 情報科学技術フォーラム講演論文集, Vol.11, No.4, pp.387-390 (2012).
- [11] 鳥居 耕太, 河中 治樹, 長瀬 功児, 小栗 宏次, ドライバ運転感覚の定量化に向けた筋電位特徴量抽出, 情報処理学会研究報告高度交通システムとスマートコミュニティ (ITS), Vol.2017-ITS-70, No.7, pp.1-6 (2017).
- [12] 小林 一夢, 秋月 拓磨, 高橋 弘毅, 章 忠, 加速度センサを用いた交差点左折時における運転操作の個人特徴に関する検討, 第 61 回自動制御連合講演会, Vol.12C3, pp.887-892 (2018).
- [13] 林 政行, 伊藤 孝行：スマートフォンを用いた走行データの収集とドライバー行動モデルの構築, 情報処理学会第 80 回全国大会講演論文集, Vol.2018, No.1, pp.403-404 (2018).
- [14] Chalermopol, P., Saiprasert, C., Pholprasit, T.: A Method for Driving Event Detection Using SAX with Resource Usage Exploration on Smartphone Platform, EURASIP Journal on Wireless Communications and Networking, Vol.2014, No.1, pp.1-11 (2014).