

誤り訂正学習による動作認識の試み

坂野 鋭^{†1} 山田 武士^{†1} 石黒 勝彦^{†1}

機械操作行動の認識について初期的な結果を得たので報告する。ロボット、自動車など人間と共生する知的機械が人間の動作を認識することは、コミュニケーション、安全確保、利便性向上の観点から重要である。本論文では自動車運転者の操作行動を認識する問題を取り上げる。操作行動認識は安全確保、利便性の観点から多くの可能性がある。本論文ではこの問題に著者が提案している「パターン認識の初期検討プロセス」を適用することでスムーズに問題を解決できた。また、Self-training による運転者個人適応についてもあわせて検討したので報告する。

A Preliminary Study of Human Motion Recognition Using Error Correction Learning

HITOSHI SAKANO,^{†1} TAKESHI YAMADA^{†1}
and KATSUHIKO ISHIGURO^{†1}

In this paper we report a preliminary study of human operation recognition in a car driving environment. Operation recognition have various potential application for improving safety and usability. We applied “preliminary process for pattern recognition study” for this problem. We easily solve this problem. We also study personal adaptation for driver operations. This problem is a new class of classification problem that includes both labeled and unlabeled categories. For solving this problem, we employ Learning Vector Quantization and self-training algorithm. Our experimental results show that the proposed method is a first step to solve this problem.

^{†1} NTT コミュニケーション科学基礎研究所

NTT Communication Science Laboratories, NTT Corporation

1. はじめに

日常生活の中で人間と共生する知的機械の実現にあたり、知的機械による人間行動の理解は、コミュニケーション、利便性向上、安全性確保など様々な観点で重要である¹⁾。このことは挨拶が理解できるロボットといった比較的遠い未来のことではなく、すでに日常的に存在する家電や自動車などの乗り物のような機械についてもあてはまる。たとえば自動車の運転に関していえば、ミラーの調整のために手を伸ばす行動を予測して手元ランプをつける、いつまでもコンソールを触っているようであればよそ見運転を警戒して警告をあげる、など多様な応用が考えられる。

特に自動車運転者の行動については交通安全の要請が強いことからカメラ画像による頭部の姿勢推定²⁾、動画像からの瞬きの検出³⁾などによる、よそ見運転、居眠り運転防止などの研究が活発に行われ、一部は実用に供されている。このため、すでに一部の乗用車には運転者を撮影するためのカメラが搭載されている。

本論文では、こうした自動車運転者の事故防止および、利便性向上の観点から車載カメラによる運転者動作の認識の問題を取り上げる。先にも述べたとおり、運転者がある機器を必要以上の時間操作し続けていればよそ見運転が発生している可能性が高い。また、このとき視線のみが機器を向いている場合も十分考えられるため、頭部姿勢推定のみではよそ見運転の完全な排除は不可能である。動作認識を併用することで安全性を高めることができると考えられる。また、これからの動作を予測することができれば、機器をアイドル状態からアクティブな状態に切り替えることも可能である。さらに、機器によっては待機電力の節約などの効果もあるかもしれない。

通常、人物動作の認識はジェスチャ認識の問題ととらえられ、隠れマルコフモデルなどの時系列データ分類の問題として解かれることが多い⁴⁾。しかし、問題をよく見てみるとジェスチャ認識というより運転者の姿勢の認識の問題に近いことが分かる。たとえば、ある機器を操作し続けていることを認識する問題は、車内の固定された運転者が固定された位置の機器に左手を置いている姿勢であり、動作ではない。また、手がステアリングを離れて車載機器に向かう動作は、通常1秒以内で終了するため、認識後のサービスのためにシステムが必要とする時間を考えると0.1秒程度、すなわち3フレーム程度の画像しか用いることができない。つまり、時系列的な方法を直接的に適用することは難しく、人体特徴点から各フレームごとに姿勢を認識する問題にむしろ近いと考えられる。

こうした観点から我々は運転者動作認識問題を人体特徴点による識別問題ととらえて検討

を開始した。新しい問題に取り組むにあたって、著者の1人が以前から提案している、「パターン認識の初期検討プロセス」⁵⁾に忠実に従い検討を進めたところ、運転者の動作を事前に学習することが可能な場合についてはKohonenの学習ベクトル量子化が有効であることを早期に判定し、非常にスムーズに問題を解決することができた。

しかし、実験の結果から動作を学習した運転者と認識対象の運転者が異なる場合には認識率の大きな低下があることが明らかになった。つまり、新たに乗用車を購入するなどした運転者の動作の認識は困難ということになる。この問題は運転者個人の動作を逐次的に学習することで解決される。そこで半教師あり学習の手法の1つであるSelf-trainingを適用したところ部分的に問題を解決することができた。

本論文では、我々が運転者動作認識という新しい認識問題に挑戦するにあたっての初期の検討経緯を事例として紹介する。以下、§2ではパターン認識の初期検討プロセスを概説し、§3ではそれに従い車載機器の操作行動認識の実験的検討を報告する。§4ではその結果を受けて個人適応の方法としてのSelf-Training法を適用した結果を報告する。§5はまとめと今後の課題である。

2. パターン認識の初期検討プロセス

画像認識に限らずパターン認識技術の検討が困難なのは学習・識別機械が数十次元から時として数万次元という不可視な高次元の特徴空間で動作するため、現実の動作が観察できず、何らかの現象が観測されても、その原因を特定することが困難であるという理由による。そのため、不慣れな研究者はそのときの論文に最もよく用いられる、すなわち制御の方法や利用上のノウハウがよく分かっていない識別器を用いるという誤りに陥りやすい。

我々が提案する初期検討プロセスはこうした誤りを避け、それが可能な場合には簡単な方法での問題解決を導くために設計されたものである⁵⁾。実のところ実用が検討されるような識別問題はFisherの判別分析のような簡単な方法で十分な認識率が得られる場合が多い。

なお、この初期検討プロセスは浜本¹¹⁾、Jainら¹²⁾が推奨している「簡単な識別器からテストする」というポリシー、および梅田¹³⁾の報告に代表される、複雑な識別器であっても良好な特徴を入力すると性能が上がるといった現象などをふまえ、著者の経験から具体化したものである。以下、順次ステップを追って解説する。

目標設定 パターン認識問題の場合通常は認識率100%を達成することは難しいか事実上不可能である。これは多くの場合、「パターン」というものが定義を持たず、何が正解であるかが判然としない場合が多いという原理的な問題に起因する。またアプリケーション側の要

求によっては必ずしも高い認識率が要求されない場合もある。実際、従来完全に人手で行っていたものを50%程度自動化できれば十分に役立つというような応用も存在する(たとえば文献6))。また適用現場の制約から処理量が小さいことを要求されることも多い。こうしたアプリケーションの要請から適切な目標認識率や要求される処理量を策定する。アプリケーションが明確でない場合には同じ分野の研究の既存研究を参考にしたり、アプリケーションの要請が明確になるのを待ったりして段階的な目標設定を行う。

データベース設計 実験に用いるためのデータを収集し、正解ラベルを付与するなどの整備を行う。このときの注意事項は、1)できるだけ大量のデータを収集する、2)アプリケーションの要請と矛盾しないデータを収集する、3)カテゴリによるデータ数の偏りなどデータの統計に異常がないかを調べておく、などである。特に注意しなくてはならないのは2)のアプリケーションの要請と矛盾しないことである。極端な例であるが「顔向き変動に頑健な顔画像認識手法」の研究のために正面顔ばかりを集めたデータベースを整備しても意味がないことは明らかであろう。ここまで極端な例は少ないがよく見ると研究の目的とデータベースの設計が整合していない例は散見される。

前処理と特徴抽出 これらの処理はアプリケーション依存性が高いため、一般論以上のことは述べないが既存文献の調査などから十分な調査を行うことが必要である。一方でデータマイニングの問題のように数値が顕わに与えられる場合には省略されることもある⁷⁾。

可視化 前に述べたとおりパターン認識機械は高次元空間で動作するために、その実態の把握は困難である。そのため、何らかの方法で2-3次元の空間に写像して、カテゴリごとの分布の様子を観察することは重要である。こうした次元圧縮の方法には主成分分析をはじめ判別分析やクラスタ判別法など様々な方法が存在する⁸⁾。できるだけ多様な方法で可視化を試み、分布やカテゴリ間の関係を把握することが必要である。むしろ、高次元空間から低次元空間に写像することでかなりの情報が落ちてしまうが重要なヒントが得られることは多い。実際、この段階でカテゴリが明確に分離していることが明らかになる場合も多い。また、可視化の結果から特徴抽出に立ち返った検討を行うことも有効である。異なる特徴抽出系の可視化結果の比較により特徴抽出系の設計に役立つ場合もある。

初期認識実験 識別器には大きく分けてカテゴリごとに単一のGauss分布を仮定した識別器と複雑な分布を仮定した識別器がある。一般的にGauss分布の仮定が正しい場合には前者の方が高い識別精度を達成する場合が多い¹⁾。また、処理量も少なく、カテゴリ追加が容易であるなどメリットが多い。後者については、確かに複雑な分布をしている場合に高い認識性能が得られるが、パラメータのチューニングにノウハウが必要であったり処理量が極度

に多かたりするなどの問題点を内包している場合も珍しくない。また、カテゴリごとの分布が複雑でも個々のカテゴリが十分遠く離れている場合には前者でかまわない場合もある。

これを適切に区別するためには、Gaussian を仮定した識別器として L_2 識別器、すなわちカテゴリごとの平均からのユークリッド距離を類似尺度として認識を実行する識別器と複雑な分布を仮定した識別器として 1-NN 識別器による実験を行い、実験結果を比較するのが良い。

これらの実験結果から、有用な特徴抽出系、識別器などを決定する。たとえば、 L_2 識別器の方が認識率が高い場合には、Gaussian の仮定が正しいと考え、より高度な部分空間法、マハラノビス距離などを検討する⁹⁾。明らかに 1-NN 識別器の方が精度が高い場合には、混合ガウス分布識別器 (Gaussian Mixture Classifier, 以下 GMM)、多層パーセプトロン (Multi-Layered Perceptron, 以下 MLP)、非線形 Support Vector Machine (以下, SVM) などの複雑な分布に対応した識別器の適用を検討する¹⁰⁾。このとき可視化結果を見てパラメータにあたりをつけておくことも重要である。微妙なのが 1-NN 識別器の方が L_2 識別器より「やや」認識率が高い場合である。このときには複雑な識別器の検討を始めたくるが実際には Gaussian の仮定が満たされている場合が多い。部分空間法などの分散共分散行列の違いを考慮した識別器により実験を行い、それでも 1-NN の方が優れている場合には複雑な識別器の利用を検討するべきである。最後に、 L_2 、1-NN の双方とも識別率がカテゴリ数分の 1 に近い、すなわち乱数的にカテゴリをあてはめたとときと同様な結果が出た場合には、特徴抽出系や前処理に立ち戻って検討するべきである。

以上のようなプロセスをふまえることで、処理量、パラメータ数などの点でオーバスペックな識別器を適用してしまうなどの危険を避けることができる。次節以降は具体的な問題、運転者の動作識別問題を議論する。

3. 運転者操作認識の検討

§1 でも述べたとおり、運転者の動作識別は運転者の利便性の向上、安全性の向上の双方の観点から重要である。動作には様々なものがあるが、ここでは初期検討であることもあ

*1 Gaussian の仮定が正しいかどうかは正規性検定のような方法を用いればよいと考えられる。しかし、現状では多変量の正規性検定は変数の skewness や kurtosis の計算が必要で簡便には実施できない。また、個々のカテゴリの分散に対して平均値が十分離れている場合には Gaussian の仮定が満たされていなくても、識別器として十分に機能する。つまり、ここでいう Gaussian の仮定とは必ずしも厳密に Gauss 分布に従っていることを意味するものではなく、Gauss 分布を仮定した識別器で十分な認識精度が得られる場合のことをいっている。こうした観点から正規性検定のような方法より、認識実験から判定する方が合理的であると考えている。

り、車載機器の操作行動に限定する。具体的にはステアリング、ナビ、コンソールボックス、エアコン、エアコン吹き出し口、ナビゲータシート、バックミラー、ギアの 9 種類の操作に対応する停留行動とステアリングからそれぞれの停留行動への遷移行動 8 種類の合わせて 17 カテゴリの動作識別問題とする。これらの動作は文献 14) で設定されたものと同様である。具体的な動作のスナップショットについては文献 14) を参照されたい。

通常、動作識別問題に対しては隠れマルコフモデルなどの時系列モデルが用いられるが、詳細にデータを見てみると、遷移行動は 0.3 秒程度、すなわち通常のフレームレートでは 10 フレーム程度の長さしかない。これでは時系列モデルの作成データとしては短すぎると判断し、今回の初期検討の段階では座標位置などを特徴とした識別問題として扱うこととした。以下、初期検討プロセスに従ってこの問題を検討する。

3.1 目標設定

本節の実験は最初期の検討となるため、運転者の行動は事前に完全に学習可能であることを仮定し、同一の運転者の異なるセッションにおける行動を認識することを目標とする。本検討で対象としている運転者操作データは現段階では具体的なアプリケーションをターゲットにしたものではない。そこで、初期段階の検討であることから目標認識率を 90% と設定した。本論文の段階では根拠・具体性ともに薄い目標値であるが、アプリケーションが具体化していくにつれ具体的な根拠を持った値になっていく予定である。

3.2 データベース設計

実験に用いたデータは (株) デンソー基礎研究所に設置された実際の自動車運転席を模擬したシミュレータ上で採取された。被験者の両肩、両肘、両手首の 6 カ所にマーカを設置し、乗用車に設置されているセキュリティカメラと同等なカメラ、すなわちフロントガラス上部中央付近に設置された魚眼カメラで操作行動を撮影し、画像中のマーカ x, y 座標を抽出して、正解ラベルを付与した。結果的にマーカ 6 カ所 \times 2 次元 = 12 次元の時系列座標データとなる。一連の動作を連続して行ったひと組のデータをセッションと呼ぶ。データは 6 名の運転者についてそれぞれ 47 セッションから 187 セッション採取され、最初の 40 セッションを学習データ、それ以降をテストデータとして用いた。

3.3 前処理と特徴抽出

今回用いたデータの採取にあたってはマーカを用いていることから、ノイズは非常に小さいと考えられるため、特に前処理を行わなかった。ただし、ステアリング周辺では動作の識別が不可能であることが明らかであるため、遷移動作の開始から 25% のフレームを手動で削除した。

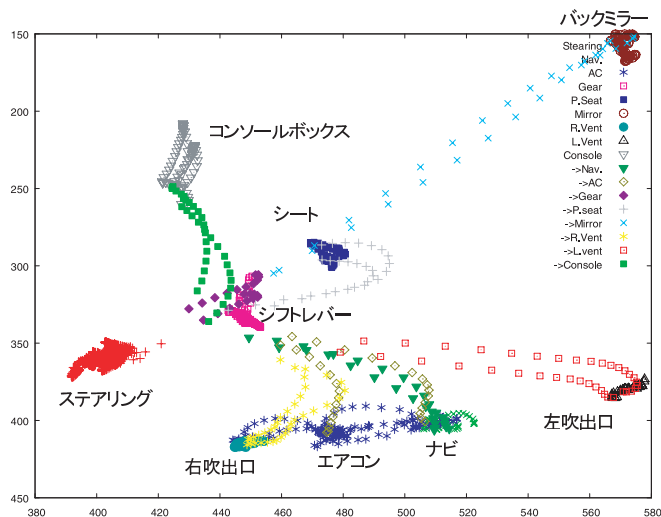


図 1 運転者の左手首動作プロット. 記号“->”は対応する停留動作への遷移動作を示す, 以下同様
 Fig. 1 Temporal plot of left wrist motions. Symbol “->” show transitions to correspond stationary states.

一般に, ジェスチャ認識の問題では座標位置や速度が特徴として用いられるが, 各動作で速度の差が明確に見られなかったことから今回の検討では座標のみを用いることにした.

3.4 可視化

可視化手法には様々な種類があるが, 今回は比較的次元であることと個々の数値の意味が明確であることから左手首, 左肘, 左肩について, それぞれ座標をプロットした. 学習データすべてのついてプロットすると軌跡がつぶれてしまうことから 1 人の運転者について 2 セッションのみをプロットした. 可視化結果を図 1, 図 2, 図 3 に示す. それぞれ手首, 肘, 肩に対応している. これらのプロットからいくつかのことが分かる.

- 図 1, 図 2 の手首, 肘のプロットでは動作ごとに比較的良好に分離している様子が分かる. 一方で図 3 の肩のプロットでは一部の動作以外はほとんど重なっており, 大きな動きの場合についてのみ識別情報を含むと想像される.
- 手首, 肘の車載機器周辺の停留動作は固定しており, よく分離している. これらについては一部を除いて Gaussian の仮定が成立しているものと推定される. これは, 車載機器の位置がカメラに対して固定されているために妥当な想定と考えられる.

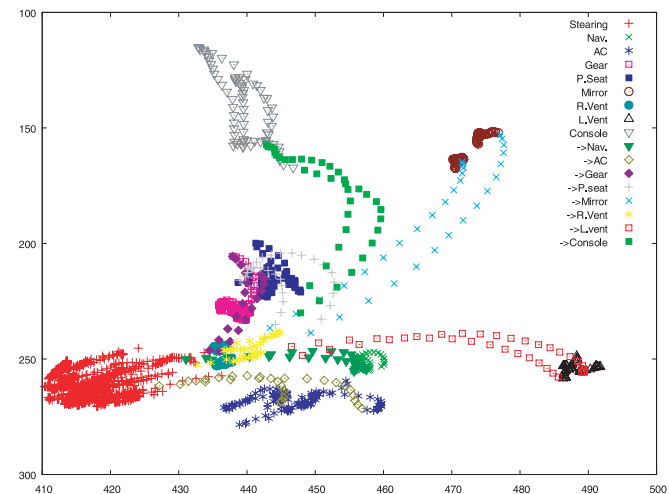


図 2 運転者の左肘動作プロット
 Fig. 2 Temporal plot of left elbow motions.

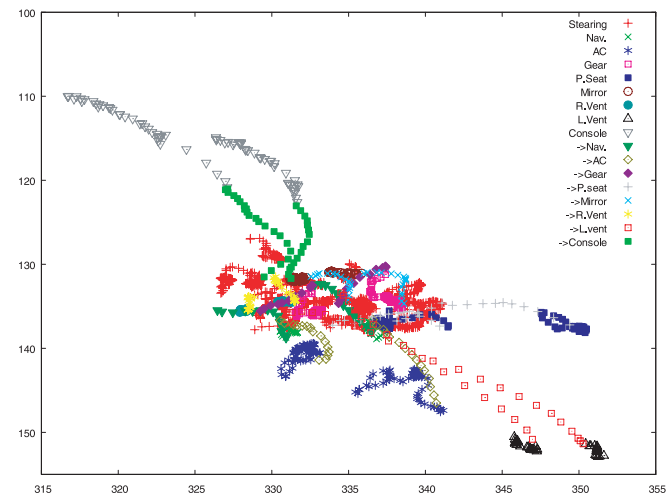


図 3 運転者の左肩動作プロット
 Fig. 3 Temporal plot of left shoulder motions.

- 手首, 肘の遷移行動については湾曲した軌跡が絡まりあっていることから Gaussian の仮定は成立していないものと考えられる.

3.5 初期認識実験

初期検討プロセスに従い, 1-NN 識別器と L_2 識別器の識別精度の比較を行った. 詳細は記述しないが 1-NN 識別器の識別精度は運転者により多少の増減はあるものの 93%程度, L_2 識別器では運転者によらず 22%程度であった.

3.6 初期実験結果に対する考察

可視化, および初期認識実験の結果から運転者動作の特徴点座標は複雑に入り組んだ分布をしており, 認識のためには Gaussian を仮定した識別器や線形 SVM, 単純パーセプトロンのように線形識別面しかつけれない識別器では不十分で, 複雑な識別面を仮定した識別器を用いる必要があることが分かった. 複雑な識別面を記述できる識別器として代表的なものが, 非線形 SVM, MLP であるが, 双方とも, 1) 学習時の処理量が多く車載システムでの個人適応学習に無理がある, 2) 非線形 SVM ではカーネルパラメータ, ソフトマージンパラメータ, MLP では中間ユニット数などのクリティカルなパラメータがあり, 高い精度を達成するためには詳細なチューニングが欠かせない, などの問題のため採用することはできない.

そこで我々は複雑な識別面が記述でき, 学習が高速であることが知られている誤り訂正学習アルゴリズム, 学習ベクトル量子化 (Learning Vector Quantization, 以下 LVQ) を採用することにした¹⁵⁾.

3.7 学習ベクトル量子化

LVQ は Kohonen により提案された学習・識別アルゴリズムであり, 各カテゴリに複数のテンプレートを持つことで複雑な識別面の記述を可能にする. LVQ には様々な改良手法が存在するがここでは最も簡単な LVQ1 を採用することにする. LVQ1 は逐次的な学習アルゴリズムであり, 識別の試行ごとにその結果に基づいてテンプレートを更新することで学習を行う. 具体的には;

$$\begin{aligned} \vec{\mu}(t+1) &= \vec{\mu}(t) - \alpha(t)(\vec{x}'(t) - \vec{\mu}(t)) & \text{if } x' \in \mathcal{C} \\ \vec{\mu}(t+1) &= \vec{\mu}(t) + \alpha(t)(\vec{x}'(t) - \vec{\mu}(t)) & \text{if } x' \in \mathcal{F} \end{aligned} \quad (1)$$

である. ここで $\vec{\mu}(t)$ は当該カテゴリのテンプレート, \mathcal{C} は正解カテゴリ, \mathcal{F} は誤読カテゴリ, \vec{x}' は認識対象の特徴ベクトル, t は学習回数, $\alpha(t)$ は収束を制御する減少関数で $0 < \alpha(t) < 1$ である. $\alpha(t)$ としては

$$\alpha(t) = \frac{\alpha_0(T-t)}{T},$$

のような関数がよく用いられる. ここで α_0 は $0 < \alpha_0 < 1$ なる定数であり, T はトータルの学習回数ですべての学習データを 1 回提示した回数である. 識別時にはそれぞれのテンプレートと識別対象サンプルのユークリッド距離を計算し, 最短となったテンプレートの属するカテゴリを認識結果として出力する.

LVQ の利点は識別時にカーネルなどの非線形関数を用いないので処理が早いことのほかに α_0 やテンプレート数などのパラメータにさほど敏感に反応しないということである. むろん, 認識率の増減はあるが SVM や MLP のようにパラメータを誤るとまったく認識ができなくなるというようなことはない. また学習時の処理量も小さいことが知られている.

3.8 認識実験

LVQ を用いて各運転者の動作を学習しているという前提の操作行動認識実験を行った. あまりチューニングができないという前提からパラメータは文献 15) に推奨値として提示されている $\alpha_0 = 0.3$, $T = 30$ を採用した. 初期値としては 40 の学習セッションのうちから 1-NN 法で他の学習サンプルを認識した結果, 最も高い認識率を達成したセッションのデータをそのまま用いた.

表 1 に認識率を示す. 後の節での運転者適応の検討のために異なる運転者に対する認識率も併記した. 表の行がテンプレートの, 列がテストデータの運転者 id である. 運転者と認識対象データが同じ, 表の対角成分ではおおむね認識率は高く, 平均して 94.8% になった. 参考のために他の代表的な手法の認識率を表 2 に示す^{*1}. このうち線形識別器, CLAFIC, 2 次識別器, 判別分析は Gaussian を仮定した識別器でそれぞれ, 共分散行列が無相関, 一

表 1 LVQ による動作認識率 (%)
Table 1 Operation recognition rates by LVQ (%).

運転者 id	1	2	3	4	5	6
1	94.7	90.6	85.4	90.4	89.6	92.4
2	89.2	96.4	70.9	84.4	87.3	90.0
3	87.2	86.9	92.8	90.3	90.4	92.6
4	89.5	89.0	83.2	94.6	88.2	91.5
5	88.0	86.1	82.9	90.0	95.6	90.7
6	87.5	92.1	78.7	80.0	85.6	94.4

*1 これらの手法の詳細は文献 9), 10), 17) などのパターン認識の標準的な教科書を参照されたい.

表 2 他の手法による動作認識率 (%)
Table 2 Operation recognition rates by other classifiers (%)

運転者 id	1	2	3	4	5	6
線形識別器 (ldc.m)	91.9	93.3	91.8	58.3	93.5	93.3
CLAFIC (subsc.m)	66.6	66.2	43.2	42.2	77.9	61.5
2 次識別器 (qdc.m)	94.0	95.1	73.1	39.5	95.0	94.2
判別分析 (fisherc.m)	74.0	80.1	79.8	23.8	84.0	80.6
GMM (混合数 10)(mogc.m)	81.2	89.1	75.6	81.7	83.7	88.6
MLP (中間層 4 素子)(bpxnc.m)	64.4	5.2	63.8	59.8	5.8	2.1
線形 SVM	87.9	87.9	86.5	51.6	93.2	57.5

部に相関がある, すべてに相関がある, すべてのカテゴリの共分散行列が等しいという仮定で導かれる. これらの結果は線形 SVM は Tübingen 大学で提供している UniverSVM¹⁶⁾ を用い, 線形 SVM 以外は Delft 大が提供している Matlab の Toolbox, PRTools4¹⁷⁾ による. PRTools4 によるものは表中に使用した Matlab スクリプト名をあげた.

これらの結果から LVQ はおおむね他手法を超える認識率を達成していることが分かる. 比較した手法のうち MLP はチューニングすれば LVQ を超える認識率を出す可能性はあるが今回は応用上の要請から行っていない.

4. 運転者個人適応の検討

特定の運転者の動作を事前に学習した場合には LVQ で良好な操作行動認識が可能であることが分かった. しかし, 表 1 の非対角成分, つまり他人の動作を学習した結果を用いて動作の認識を行った場合を見ると著しい認識率の低下が起こることが分かる.

つまり, 新車を購入した運転者などの新しいユーザの動作を認識することは, 出荷前の学習結果を用いるだけでは困難であり, 何らかの再学習が必要であることを示している. しかし, 新しい運転者に教師あり学習を適用するためには, 機械の指示に従って複数回の動作を行うことを強要することになり現実的ではない. したがって, 運転者のふだんの動作から徐々に適応していくアルゴリズムの開発が必要になる.

今回のように動作の個人性に適応する問題はジェスチャ認識の研究では筆者らの知る限り見られない. 周辺分野で活発に研究が行われているのは音声認識における話者適応であるが, この分野では基本的に指定したテキストを読み上げることにより適応を行っている技術が大半である¹⁸⁾.

これに対し, 運転者適応の問題が特異なのはステアリング, ナビ, エアコン, ギア, コン

ソールボックスでの停留行動およびそこに至る遷移行動では機器を操作した行動から電気信号の形で教師信号を受け取ることができるのに対し, それ以外の助手席シート, バックミラー, 左右吹き出し口での停留行動およびそこに至る遷移行動では教師信号が与えられないという点である.

この問題を解く試みとして我々は Self-training として知られている枠組みを用いた¹⁹⁾. Self-training は McLachlan により正準判別分析に対する半教師付き学習法として提案され, 近年の半教師付き学習研究の中で再発見された方法である²⁰⁾. Self-training のアルゴリズムは, 1) ラベル付きのデータのみを用いて識別器を学習する, 2) この識別器を用いてラベルなしのデータに仮のラベルを与える, 3) 仮のラベルを用いて識別器を学習する, 4) 2) に戻る, という簡便なもので, 理論的な根拠は希薄なものの音声認識²¹⁾, 文字認識²²⁾, 物体認識²³⁾, テキスト分類²⁰⁾ などの分野で実験的に良好な結果が得られている*1. この方法はどのような学習・識別系にも適用可能であるというメリットがある. このため今回の初期的な検討のためには適当であると考えた.

Self-training は学習前の識別率が高くないと良好な学習結果が得られないという性質がある. また, McLachlan の解析結果から初期の学習しか効果がない, つまりいたずらに学習を続けても認識率の向上が見られないことも知られている. そこで表 1 の結果から他の運転者の動作の認識率が最も高いテンプレートを初期値として実験を行った. 当初, Kohonen の推奨値のとおり $T=30$ で実験を行ったが, 認識率向上が見られなかったため McLachlan の示唆に従い $T=10$ とした. 実験データについてはすべての動作にラベルがあるため, 教師のないと想定される動作についてはラベルを外して実験を行った. Self-training は自由度の高い方法なのでいくつかの適用方法が考えられる. 今回は以下の 3 種類の方法で実験した. 他の運転者の学習結果を初期値として用いた Self-training LVQ

- (1) 表 1 の認識率を用いて, 適応実験対象者の認識率が最も高い他の運転者のテンプレートを選ぶ.
- (2) 適応実験対象者の学習データのうちラベルが付与されていないデータについて選択したテンプレートによる認識処理を行い, 認識結果を仮ラベルとして付与する.
- (3) 選択したテンプレートを LVQ の初期値として, 仮ラベルを付与された適応実験対象者の学習データで LVQ の学習を行う.
- (4) 学習結果を用いて再び認識処理を行い仮ラベルをつけなおす.

*1 学習前の分布の仮定が正しければ Self-training は良好に動作するという報告もある²⁴⁾.

表 3 他人テンプレートを初期値とした, Self-training LVQ による動作認識率 (%)

Table 3 Operation recognition rates by Self-training LVQ (%) using other drivers template as initial template.

運転者 id	1	2	3	4	5	6
本人	94.7	96.4	92.8	94.6	95.6	94.4
教師あり	92.5	96.1	92.5	93.4	93.5	94.2
Self-training	91.1	92.8	88.8	93.2	92.8	94.1
学習前	89.5	92.1	85.5	90.4	90.4	92.6

表 4 運転者本人のデータから初期値を生成した, Self-training LVQ による動作認識率 (%)

Table 4 Operation recognition rates by Self-training LVQ (%). Initial templates are generated from the drivers themselves.

運転者 id	1	2	3	4	5	6
本人	94.7	96.4	92.8	94.6	95.6	94.4
Self-training	92.6	93.9	90.1	94.7	94.7	93.2
学習前	91.3	93.0	89.2	93.9	94.3	91.2

(5) 仮ラベルを付与したデータで再学習する。以降これを繰り返す。

運転者本人のデータから初期値を生成した Self-training LVQ

- (1) 表 1 の認識率を用いて, 適応実験対象者の認識率が最も高い他の運転者のテンプレートを選ぶ。
- (2) 適応実験対象者の学習データのうちラベルが付与されていないデータについて選択したテンプレートによる認識処理を行い, 認識結果を仮ラベルとして付与する。
- (3) 学習データから 1 セッション分を選びこれを 1-NN 法のテンプレートとして他の学習データを認識する。この認識率が最も高かった 1 セッション分の学習データを LVQ の初期値とする。
- (4) 仮ラベルを付与された学習データで LVQ の学習を行う。
- (5) 学習結果を用いて再び認識処理を行い仮ラベルをつけなおす。
- (6) 仮ラベルを付与したデータで再学習する。以降これを繰り返す。

Self-training k-NN

- (1) 表 1 と同様に 1-NN 法での認識実験を行い, 適応実験対象者の認識率が最も高い他の運転者を選ぶ。
- (2) 他の運転者の学習データを用いて適応実験対象者の学習データに仮ラベルを付与する。

表 5 Self-training 1-NN による動作認識率 (%)

Table 5 Operation recognition rates by Self-training 1-NN (%).

運転者 id	1	2	3	4	5	6
本人	92.7	95.6	90.3	93.4	95.5	91.9
Self-training	92.6	91.3	85.6	93.5	95.4	92.4
学習前	90.0	89.9	82.4	90.0	90.6	92.0

それぞれ学習後の認識率を表 3, 表 4, 表 5 に示す。6 人の運転者すべてについて認識率が向上していることが分かる。

5. まとめと今後の課題

自動車運転者の行動を対象に動作の認識と個人適応に対する初期的な検討を行った。運転者動作の認識は動作の認識というより姿勢の認識に近いことから, 静的なパターン認識問題ととらえ, 著者が提案しているパターン認識研究の初期検討プロセスを適用して検討したところ, スムーズに問題を解決することができた。

また, 検討途上で学習した運転者と異なる運転者の動作認識率が低いという問題に遭遇し運転者動作の個人適応の問題も検討した。この問題は教師なしのカテゴリが複数存在する新しいクラスの識別問題であり, これまでに取り組みがなかったが Self-training を用いた検討により, ある程度解決が可能であることが分かった。

しかし, Self-training は初期の学習しか効果がないため, 動作データが次々に与えられても認識率が向上しないという明らかな限界があり, この問題の最終的な解決にはなっていない。また, Self-training の初期値を単純に他の運転者動作の認識率が高いものという基準で選択しているが, これは応用からすると実行可能な操作であるか疑問である。何らかの基準で運転者の類似度を評価する仕組みが必要である。

今後はより多数の運転者を対象に上記の問題の解決を図るための検討を行う。また, 運転操作以外の動作についても検討を広げる予定である。

謝辞 運転者行動データの提供および使用の許諾をしていただいた, 株式会社デンソー基礎研究所福本晴継第一研究室室長, 伊藤隆文主任研究員, 石川貴洋研究員に感謝いたします。

参考文献

- 1) Pentland, A.: Looking at People: Sensing for Ubiquitous and Wearable Computing, *IEEE Trans. PAMI*, Vol.22, No.1, pp.107-119 (2000).

- 2) 高橋大介, 岡本教佳: 明るさの変化に口バスタな運転者の姿勢推定, *電学論 C*, Vol.127, pp.1043-1051 (2007).
- 3) 足立和正, 浜田尊裕, 中野倫明, 山本 新: ドライバの意識低下検知のための動画像処理によるまばたき計測, *電学論 C*, Vol.124, pp.776-783 (2004).
- 4) Yamato, J., Ohya, J. and Ishii, K.: Recognizing Human Action in Time-Sequential Images using Hidden Markov Model, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (1992).
- 5) 坂野 鋭, 山田敬嗣: 怪奇!! 次元の呪い—識別問題, パターン認識, データマイニングの初心者のために, *情報処理*, Vol.43, No.5, pp.562-567, No.6, pp.658-663 (2002).
- 6) 東 陽子, 佐藤 新, 樋口裕高, 坂野 鋭, 森本俊彦, 松永 務, 石井敬介, 村松 正明: クラスタリングによる遺伝子解析—遺伝子型の自動決定, *信学技報 PRMU2002-71* (2002).
- 7) 坂野 鋭: データマイニングの脈脈を探る—パターン認識の視点から, *信学技報 PRMU2003-93* (2003).
- 8) 末永高志, 佐藤 新, 坂野 鋭: クラスタ構造に着目した特徴空間の可視化法—クラスタ判別法, *信学論 D-II*, Vol.J85-D-II, pp.785-795 (2002).
- 9) 石井健一郎, 上田修功, 前田英作, 村瀬 洋: わかりやすいパターン認識, オーム社 (1998).
- 10) ビショップ, C.M.: *パターン認識と学習機械*, シュプリンガー・ジャパン (2008).
- 11) 浜本義彦: 統計的パターン認識: 過去・現在・未来, *信学技報*, PRMU2000-129 (2000).
- 12) Jain, A.K., Duin, R.P.W. and Mao, J.: Statistical Pattern Recognition: A Review, *IEEE Trans. PAMI*, Vol.22, pp.4-37 (2000).
- 13) 梅田三千雄: 手書き漢字認識手法の個性的なサーベイ, 電子情報通信学会「手書き文字認識技術の過去・現在・未来」シンポジウム講演論文集, pp.54-63 (1993).
- 14) Ito, T. and Kanade, T.: Predicting Driver Operations inside Vehicles, *Proc. IEEE Intl. Conf. on Automatic Face and Gesture Recognition FG2008* (2008).
- 15) T. コホネン: 自己組織化マップ, シュプリンガー・フェアラーク東京 (2005).
- 16) <http://www.kyb.mpg.de/bs/people/fabee/universvm.html>
- 17) <http://prtools.org/>
van der Heijden, F., Duin, R.P.W., de Ridder, D. and Tax, D.M.J.: Classification, Parameter Estimation and State Estimation, Willy (2004) に PRTools で扱える識別器などについての解説がある.
- 18) 篠田浩一: 確率モデルによる音声認識のための話者適応化技術, *信学論 D-II*, Vol.J-27-D-II, pp.371-386 (2004).
- 19) McLachlan, G.J.: Iterative reclassification procedure for constructing an asymptotically optimal rule of allocation in discriminant analysis, *J. Amer. Statist. Assoc.*, Vol.70, pp.365-369 (1975).
- 20) Chapelle, O., Schölkopf, B. and Zien, A. (Eds.): *Semi-Supervised Learning*, MIT

- press (2006).
- 21) Kemp, T. and Waibel, A.: Unsupervised Training of a Speech Recognizer: Recent Experiments, *Proc. Eurospeech*, Vol.6, pp.2725-2728 (1999).
- 22) Nagy, G.: Classifiers That Improve with Use, *信学技報*, PRMU2003-228 (2004).
- 23) Rosenberg, C., Hebert, M. and Schneiderman, H.: Semi-supervised self-training of object detection models, *7th IEEE Workshop on Applications of Computer Vision* (2005).
- 24) Kuncheva, L.I., Whitaker, C.J. and Narasimhamurthy, A.: A case-study on naive labelling for the nearest mean and the linear discriminant classifiers, *Pattern Recognition*, Vol.41, pp.3010-3020 (2008).

(平成 21 年 2 月 5 日受付)

(平成 21 年 3 月 27 日再受付)

(平成 21 年 4 月 22 日採録)



坂野 鋭 (正会員)

1988 年中央大学理工学部物理学科卒業。1990 年埼玉大学大学院理学研究科修士課程修了。同年 NTT データ通信(株)(現(株)NTT データ)入社。2008 年より NTT コミュニケーション科学基礎研究所主任研究員。統計的パターン認識と学習理論の研究に従事。博士(工学)。1998 年情報文化学会論文賞受賞。日本物理学会, 電子情報通信学会, 画像電子学会,

IEEE 各会員。



山田 武士 (正会員)

1988 年 3 月東京大学理学部数学科卒業。同年 NTT 入社。1996 年より 1 年間英国コベントリー大学客員研究員。NTT コミュニケーション科学基礎研究所知能創発環境研究グループリーダー等を経て, 現在, NTT 先端技術総合研究所研究推進担当部長。主として機械学習, メタヒューリスティクスによる組合せ最適化等の研究に従事。博士(情報学)。電子情報通信

学会, ACM, IEEE 各会員。



石黒 勝彦

2006年東京大学大学院情報理工学系研究科修士課程修了。同年NTTコミュニケーション科学基礎研究所。機械学習，パターン認識，時系列解析の研究に従事。電子情報通信学会，IEEE各会員。
