

誤り訂正学習による動作認識と個人適応への試み

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し、個人の行動の個性に適応していく問題は単なる学習問題ではなく、複数の教師あり学習と教師無し学習が混在する新しいクラスの問題である。本論文ではこの問題に対して特に自動車の操作行動を対象として初期的な検討を行った。学習ベクトル量子化と Self-training を適用することにより、これらの問題への解決可能性を示す。

A Preliminary Study of Human Motion Recognition and Personal Adaptation Using Error Correction Learning

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

In this paper we report a preliminary study of human motion recognition and its personal adaptation in a car driving environment. 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. はじめに

日常生活の中で人間と共生する知的機械の実用化にあたり、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。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、既に日常的に存在する家電や自動車などの乗り物の様な機械についてもあてはまる。例えば自動車の運転に関して言えば、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様であればよそ見運転を警戒して警告を上げるなど、多様な応用が考えられる。

これらの研究は一般にジェスチャ認識の問題と捉えられ、その動作を学習した識別器によって解決されることが多い。しかし、問題をよく見てみると新しいクラスの学習の問題が提示されていることに気付く。

後の節で実験の対象とする自動車の車載機器操作を例に取って考えよう。車載機器の操作にはギアの操作の様な運転に関連した動作、エアコンやナビなどの車載機器の操作、ミラーや空気吹き出し口の様な調整操作がある。いずれの動作も機器の位置が固定されているために、どの運転者が操作したとしても基本的には同じような動作になると予想される反面、ハンドルの持ち位置や腕の長さの違いにより、操作行動の軌道に

は大きな違いが存在しうる。つまり、操作される側の機械が人間行動を予測するためには、標準的な動作を学習するだけでは不十分で、個々の運転者の動作を逐次学習して適応していかななくてはならないことになる。

ここで問題になるのはギアやエアコンなどの操作を行った場合、操作に伴う電気信号を教師信号として用いることで教師付き学習による適応を実現できるのに対して、ミラーや吹き出し口の操作について正解を与えることが困難であるということである。つまり、標準的な運転操作学習結果から個々の運転者に適応するためには完全に教師付きのカテゴリに関するデータと教師の無いカテゴリに関するデータが混在する状態で学習を進めなくてはならないことになる。

本論文では、この新しいクラスの学習・識別問題に対しての初期的な検討結果を報告する。著者の一人が推奨している「パターン認識の初期検討プロセス」に従って、学習識別機械として学習ベクトル量子化と Self-training を利用した適応を行ったところ、良好な結果を得た。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

^{†1} NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

実の動作が観察できないという理由による。そのため、不慣れた研究者はその時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、簡単な方法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²⁾。実のところ実用が検討される様な識別問題は Fisher の判別分析の様な簡単な方法で十分な高い認識率が得られる場合が多い。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が事実上不可能である。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様な応用も存在する。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つ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。特に注意しなくてはならないのは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無いことは明らかであろう。

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様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。

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³⁾。出来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。

初期認識実験 識別器には大きく分けてカテゴリあたりに単一の Gaussian 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。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、即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。例えば、 L_2 識別器の方が認識率が高い場合には、明らかに 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⁴⁾。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)、Support Vector Machine (以下, SVM) 等の複雑な分布に対応した識別器の適用を検討する⁵⁾。このとき可視化結果を見てパラメータあたりをつけておくことも重要である。最後に、双方ともカテゴリ数分の 1 に近い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。以上の様なプロセスを踏まえることで、過剰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。

3. 運転者操作認識の検討

運転者動作には様々なものがあるが、初期検討であることもあり、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、ナビゲーターシート、バックミラー、ギアなどへの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の計 17 カテゴリの動作識別問題とする。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。

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シミュレータ上で両肩、両肘、両手首の 6 か所にマーカーを設置し、高級車に設置されているセキュリティカメラと同等な魚眼カメラにて操作行動を撮影し、マーカー位置を抽出、正解ラベルした 12 次元の時系列座標データである。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 件を学習データ、それ以降をテストデータとして用いた。また、全ての動作はステアリングから始まるが、ステアリング周辺では識別が不可能であることが明らかであるため、動作開始から 25% については手動で削除した。これ以上の詳細な条件については文献⁶⁾ 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⁷⁾、詳細にデータを見てみると、遷移行動は 0.3 秒程度、即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、初期検討の段階では座標位置を特徴とした識別問題として扱うこととした。

3.1 目標設定

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 可視化

今回は比較的低位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。学習データ全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一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に示す。

これらのプロットからいくつかのことがわかる。

- 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わかる。一方で肩のプロットでは一部の動作以外は重なっており、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。
- 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。
- 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。

データベースの設計については前述した。また、今回は特徴抽出は行わず座標データの数値を生で用いるため、次節では初期認識実験について述べる。

3.3 初期認識実験

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、L₂識別器では運転者によらず22%程度であった。

3.4 初期実験結果に対する考察

可視化、及び初期認識実験の結果から運転者操作行動の認識のためには Gaussian を仮定した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わかった。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、SVM,MLPであるが双方とも1)処理量が多く車載システムでの利用に無理がある、2)SVM,MLPともに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出来ない。

そこで我々は複雑な識別面が記述出来、処理が軽い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⁸⁾。

3.5 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

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

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;

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

である。ここで $\bar{\mu}(t)$ 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、 \bar{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 $0 < \alpha(t) < 1$ であり、 $\alpha(t) = \frac{\alpha_0(T-t)}{T}$ の様な減少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 は全学習回数で全ての学習データを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等の非線形関数を用いないので処理が早いこと他に α_0 やテンプレート数などのパラメータにさほど敏感ではないことである。

3.6 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出来ないという前提からパラメータは文献⁹⁾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、表の対角成分では認識率は高く、平均して94.8%になった。

4. 運転者個人適応の検討

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及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出来るのに対し、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及びそこに至る遷移行動では教師信号が与えられない点である。

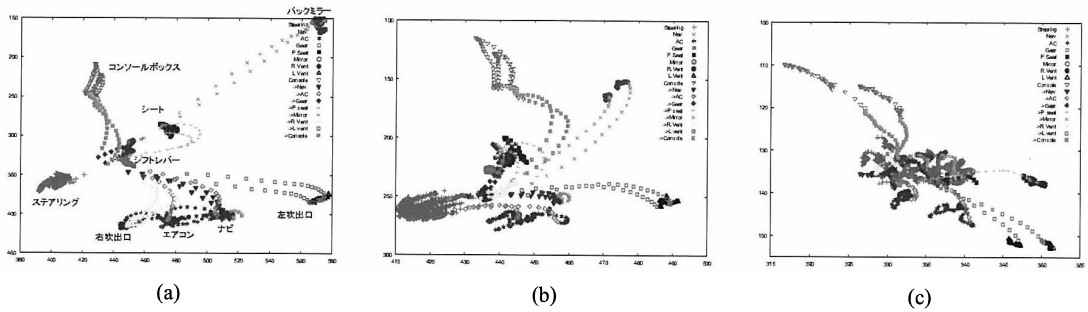


図 1 運転者の (a) 左肩, (b) 左肘, (c) 左肩動作プロット
Fig. 1 Temporal plots of left (a)wrist, (b)elbow, (c)shoulder motion.

表 2 Self-training による動作認識率 (%)
Table 2 Operation recognition rate by Self-training(%)

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

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⁹⁾。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, どの
様な学習・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 は最初の識別率が低いと良好な学
習結果が得られないという性質がある。また, McLa-
clan の解析結果から初期の学習しか効果が無いことも
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
T=10 として実験を行った。

学習後の認識率を表 2 に示す。わずかであるが認識
率が向上していることがわかる。

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。個人適応の問題は教
師無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無かったが LVQ
と Self-training を用いた検討により, ある程度解決が
可能であることがわかった。

しかし, Self-training は初期の学習しか効果が無い
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。

謝辞 運転者行動データの提供及び使用の許諾をし
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) A. Pentland, "Looking at People: Sensing for Ubiquitous and Wearable Computing," IEEE Trans. PAMI, Vol. 22, No.1, pp.107-119, (2000)
- 2) 坂野 鋭, 山田敬嗣, 怪奇!! 次元の呪いー識別問題, パターン認識, データマイニングの初心者のためー, 情報処理, Vol. 43, No. 5, pp.562-567, No.6 pp.658-663, (2002)
- 3) 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, J85-D-II, No. 5, pp.785-795. (2002)
- 4) 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)
- 5) C. M. ビショップ, パターン認識と学習機械, シュプリンガー・ジャパン, (2008)
- 6) T. Ito, T. Kanade, "Predicting Driver Operations inside Vehicles," In Proc. IEEE Intl. Conf. on Automatic Face and Gesture Recognition, (2008).
- 7) J. Yamato, J. Ohya and K. Ishii, "Recognizing Human Action in Time-Sequential Images using Hidden Markov Model," In Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (1992)
- 8) 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)
- 9) G. J. McLachlan, "Iterative reclassification procedure for coconstructing an asymptotically optimal rule of allocation in discriminant analysis," J. Amer. Statist. Assoc., Vol. 70, pp.365-369, (1975)