

特集 漫然運転検出のための適応的ブレーキ警報システムの開発*

An Adaptive Rear-End Collision Warning System for Drivers

坂東 誉司	池田 和司	味間 弘喜	井上 裕太
Takashi BANDO	Kazushi IKEDA	Hiroki MIMA	Yuta INOUE
柴田 智弘	人見 謙太郎	深谷 直樹	
Tomohiro SHIBATA	Kentaro HITOMI	Naoki FUKAYA	

This paper proposes a rear-end collision warning system, where the collision risk is adaptively set from driving signals. The system employs the inverse of the time-to-collision with constant relative acceleration as the risk and the one-class support vector machine as the anomaly detector. The system also utilizes brake sequences for outlier detection. When a brake sequence has a low likelihood with respect to trained hidden Markov models, the driving data during the sequence are removed from the training dataset. This data selection is confirmed to increase the robustness of the system by computer simulation. Consequently, the effectiveness of the proposed system is also confirmed by evaluation experiments.

Key words: Collision warning system, Time-to-collision, Support vector machine, Hidden Markov model

1. はじめに

近年交通事故の数は減少傾向にあるものの、その数は依然として大きく、交通事故を更に減らすために衝突リスク指標の検討や緊急ブレーキ支援の開発が盛んに行われている¹⁾⁷⁾。更にドライバの運転行動予測や運転状況推定も近年重要視されている⁸⁾⁹⁾。ドライバの行動や運転状況を推定できれば、衝突防止システム等の運転支援機能を運転状況に合わせて適用することができるため、事故を未然に防ぐという観点からみて有効である。例えばFCW (Forward Collision Warning) の様な衝突警報システムでは、従来、警報を出す基準があらかじめ定められていた。このため、ドライバによっては警報が過度に作動すると感じたり、逆に警報が遅いと感じたりする。通常時の運転と異常時の運転を判別し、ドライバや運転状況に合わせて警報を出す基準を設定できれば、ドライバ受容性の高い警報システムの構築につながる。そこで本論文では、車間距離に基づく衝突リスク指標を利用し、機械学習技術を利用して警報のしきい値を適応的に設定するシステムについて考える。

本論文の構成は下記の通りである。まず、2節で機械学習法の一つである One Class Support Vector Machine (OCSVM) を用いた通常運転の学習法について述べる。また、3節では OCSVM を運転データに適用する上での問題点の解決法について述べる。4節では通常運転学習をロバストに行うための、学習に用いる運転データの取捨選択法について説明し、5節の評価実験を通して提案手法の効果を確認する。最後に6節で本論文をまとめる。

2. OCSVM を利用した通常運転学習

運転挙動データを通常データと異常データに分離することは、言い換えると観測された運転挙動データについてリスク指標の適切なしきい値を設定することである。こうした通常データと異常データを分離する課題に対して機械学習と呼ばれる分野から、近年のデジタルカメラの顔検出に利用されているブースティングや Support Vector Machine (SVM)、ニューラルネットなど様々な識別器が提案されている。ただし、ドライバは普段安全に運転しようと努めており、異常時の運転挙動データを取得することは極めて困難である。このため、適切なしきい値の設定に異常データを利用できない問題があった。こうした手法の多くは、通常データと異常データの両方を用いて二つを分離するための「しきい値」を決めるため、ここで扱うような異常データの取得が困難な場合には適用できない。我々はこうした問題に対して、Support Vector Machine (SVM)¹⁰⁾ の一つである One Class SVM (OCSVM)¹¹⁾ が有効なことを確認している⁷⁾。

OCSVM は、例えば「データの 95%」というような基準が与えられたときに、密度関数のサポートを推定する手法である¹³⁾。これにより、通常データ領域に含まれない異常データを見つけることができる。OCSVM は超平面により、多くのデータが含まれる領域と少数のデータが含まれる領域を分離する。

今、観測された運転挙動データベクトルを $\{f_i\}_{i=1}^N$ とすると、OCSVM は ν を定数パラメータとして、

* 2010年7月30日 原稿受理

$$\min_{\xi, \rho} \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{\nu N} \sum_{i=1}^N \xi_i - \rho, \quad \dots\dots\dots(1)$$

$$s. t. \mathbf{w}^T \mathbf{f}_i \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad \rho \geq 0$$

を満足するようにベクトル \mathbf{w} とスラック変数 ξ_i 、バイアスパラメータ ρ を決定する。OCSVM では $\mathbf{w}^T \mathbf{f}_i \geq \rho$ で表される半空間をデータが分布する領域とみなしている。これは直感的に、データが境界面の外側（原点の反対側）に位置するような超平面のうち、原点からの距離が最大のものを境界面として採用していると解釈することができる。この時ベクトル \mathbf{w} は

$$\mathbf{w} = \sum_{i=1}^N a_i \mathbf{f}_i, \quad 0 \leq a_i \leq \frac{1}{\nu N}, \quad \sum_{i=1}^N a_i = 1 \quad \dots\dots\dots(2)$$

と表され、一般に解はスパースになる。すなわち、ごく少数の a_i だけが非ゼロになることが知られている。 $a_i > 0$ となる a_i に対応するデータ \mathbf{f}_i はサポートベクトルと呼ばれる。

先述の通り OCSVM では通常のデータ領域に含まれるデータの確率を決めることができる。ここでは 95% に設定したため、データ中の外れ値の割合は 5% となる。この割合を制御することでシステムの感度を調整することができ、この割合を大きくすれば検出感度が上昇し、小さくすれば検出感度は低下する。この調整はドライバの好みに応じて設定すればよい。なお、OCSVM の学習には LIBSVM¹²⁾ を利用した。

3. OCSVM の運転挙動データへの適用

前節で OCSVM を用いた通常運転学習の枠組みを概説したが、オリジナルの OCSVM は運転挙動データにおける通常運転学習には適用できない場合がある。

我々は、衝突リスク指標として衝突余裕時間 (Time-to-Collision, TTC) を採用する。衝突リスク指標は数多く提案されており、自動車の物理的特性に基づくものと、ドライバの心理的特性に基づくものに大別される¹⁴⁾。ここで採用した TTC は自車と先行車が等速直線運動を行うと仮定した上での衝突までの時間である。TTC は物理的な意味づけが明確である一方で、ドライバの網膜上でのオプティカルフローからも計算可能であることから、生理的特性からも正当化されるため、広く利用されてきた¹⁵⁾。しかし、実環境で車両が等速で走行することは稀である。そこで等速直線運動ではなく、等加速度運動を仮定した衝突余裕時間を考える。これはニュートン力学により二次方程式の解となるので 2 次 TTC あるいは TTC_{2nd} と呼ばれている。

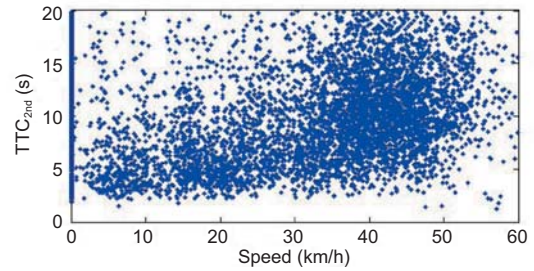


Fig. 1 TTC_{2nd} of all datasets

Fig. 1 に車速と TTC_{2nd} の分布を示す。この場合、データの存在する領域と外れ値領域を分割するためには、右上がりの判別直線が適していると思われる。ところが運転挙動データが Fig. 1 に示すように第 1 象限に存在している場合、OCSVM では式 (2) より重みベクトル \mathbf{w} が正となり、それ故に判別直線は右下がりになってしまう。我々はこの問題を解決するために、 TTC_{2nd} の代わりに TTC_{2nd} の逆数 ($ITTC_{2nd}$) を採用する (Fig. 2)。これにより、判別直線は右下がりになり、OCSVM の適用に適した形となる。

ところが $ITTC_{2nd}$ の採用により、データの大部分が判別直線の原点側に分布してしまう。先述のとおり、OCSVM では判別直線の外側（原点の反対側）のデータを通常データとして扱うため、この場合は 95% のデータを外れ値として扱わなければならない。SV の数 (OCSVM の場合は外れ値の数) に比例して計算量が増加する SVM において、これは計算量の点で大きな問題を招く。この点を解決するために、我々は Fig. 2 における原点の位置を (120, 0.3) に移動させる。高速道路での運転も含め、通常この点よりも右上にデータが存在しないことが確認されている⁷⁾。この原点の移動により、外れ値として扱うデータの数を大きく減らすことができ、高速な通常運転の判定が可能となる。

4. 学習データの取捨選択

前節で提案した警報システムでは、観測される運転挙動は全て「正常な」データであると仮定していた。ところが、

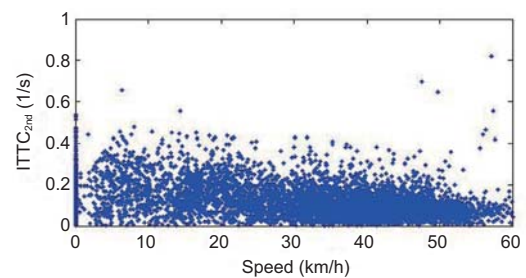


Fig. 2 $ITTC_{2nd}$ of all datasets

この仮定は常に正しいとは言えず、実際にはヒヤリハット等の異常データがノイズとして混入することになる。2節で述べた通り、SVMの解は少数のSVで構成されるが、SVになるデータは一般に境界面付近のデータとなる。このためSVMは前述のノイズの影響を受けやすく何らかの対処が必要となる。

SVM自体の工夫としてソフトマージンと呼ばれる手法も提案されているが¹⁰⁾¹⁶⁾、予め異常データを訓練データから除去する手法も有効である¹⁷⁾。ここでは、運転行動のモデル化に有効であることが知られているブレーキ系列¹⁸⁾を利用して異常データを訓練データから除去する手法を提案する。

4.1 GMMを用いたクラスタリング

運転挙動に含まれるドライバーの特性をモデル化する手法として、混合ガウスモデル(GMM)が知られている。GMMはK個のガウス分布を足し合わせた

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k N(\mathbf{x} | \mu_k, \Sigma_k) \dots \dots \dots (3)$$

で表現される。各ガウス分布 $N(\mathbf{x} | \mu_k, \Sigma_k)$ は混合要素と呼ばれ、それぞれ平均 μ_k と分散 Σ_k を持つ。また π_k は各ガウス分布の混合比を示す。ここで、K次元の2値隠れ変数 z を導入し、この z は1-of-K表現をとるとすれば、式(3)は下記のように書き直すことができる。

$$p(\mathbf{x}) = \sum_z p(z) p(\mathbf{x} | z) = \sum_{k=1}^K \pi_k N(\mathbf{x} | \mu_k, \Sigma_k) \dots \dots \dots (4)$$

ここで、 z の周辺分布 $p(z)$ 、 z を与件とする \mathbf{x} の条件付分布 $p(\mathbf{x} | z)$ は、それぞれ

$$p(z) = \prod_{k=1}^K \pi_k^{z_k} \dots \dots \dots (5)$$

$$p(\mathbf{x} | z) = N(\mathbf{x} | \mu_k, \Sigma_k)^{z_k} \dots \dots \dots (6)$$

と書ける。

GMMのパラメータ μ_k, Σ_k, π_k はEMアルゴリズム¹⁹⁾により観測データ $\{\mathbf{x}_t, t = 1, \dots, T\}$ から推定する。先に導入した z_t が、観測データ \mathbf{x}_t がどのガウス分布から生成されたかを示す指示変数になっているので、 $\sum_{k=1}^K |z_{k,t}|$ が最も大きくなるガウス関数から \mathbf{x}_t が生成されたとクラスタリングすることが可能である。

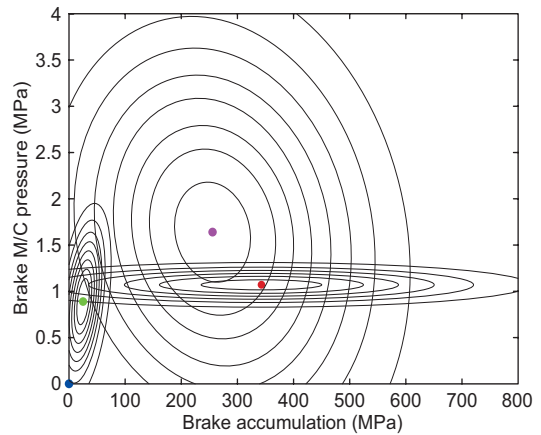


Fig. 3 Iso-probability contour of the densities of the Gaussian components

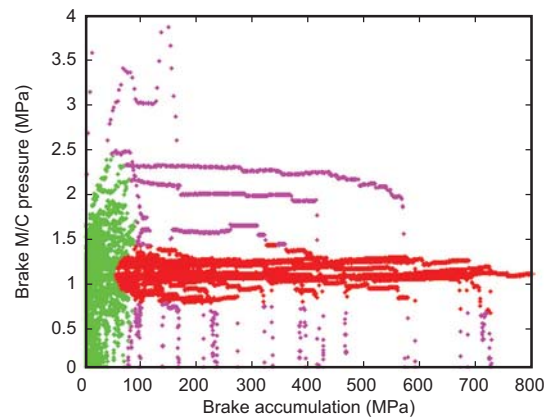


Fig. 4 Clustering of the data with the GMM

我々の実験では観測データ \mathbf{x}_t は時刻 t におけるブレーキのM/C圧と、そのブレーキを踏み始めた時点からの累積和の二次元とした。Fig. 3に26,751点の通常運転時のデータを用いて学習した、混合数4のガウス分布を示す。これによりブレーキ系列は下記の四つに分類できる。

1. 短時間のブレーキ操作
2. 長時間のブレーキ操作でM/C圧が一定値の区間が長い
3. 長時間のブレーキ操作でM/C圧が時々刻々と変化する
4. ブレーキ操作を行っていない

Fig. 4に実際のブレーキ系列の分類結果を示しておく。但し、GMMを用いたこれらの分類ではブレーキ操作が時々刻々と変化する系列であるということは考慮していない。このため、同じブレーキ系列でも時刻毎に別のクラスに分類されてしまうことがあった。次節ではこの問題に対処する手法について説明する。

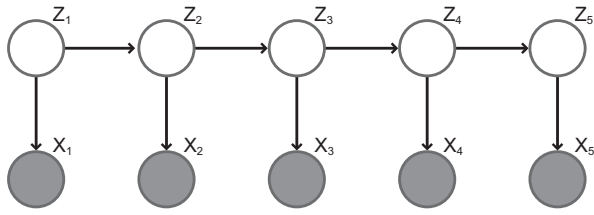


Fig. 5 Graphical model of an HMM

4.2 HMM を用いたデータ選別

隠れマルコフモデル (HMM) は音声認識, 自然言語処理, オンライン手書き文字認識などに広く用いられている系列データを扱う手法の一つであり, 系列データの伸縮 (データ長の変化) にも柔軟に対応することができる²⁰⁾. HMM では系列データを, 隠れ変数のマルコフ連鎖で表現し, 各々の観測値は対応する隠れ変数の状態によって条件付けられる観測モデルから生成されると仮定する. 前節で説明した GMM によるデータの生成モデルを, 各時刻で混合要素が独立に選択されるのでは無く, 過去の観測で選択された混合要素に依存して選択されるように, 拡張したモデルと解釈することもできる. Fig. 5 に HMM のグラフィカルモデルを示す. 変数間の依存関係が矢印で表現されており, 隠れ変数 (ここでは状態と呼ぶ) z_t は前時刻の z_{t-1} にのみ依存するモデルを仮定した. つまり,

$$p(z_z | z_{t-1}) = \prod_{k=1}^K \prod_{j=1}^K A_{kj}^{z_t, z_{t-1}} \dots\dots\dots (7)$$

と書く. ここで A_{kj} は状態 k から j への状態遷移確率を示し, $A = \{A_{kj}\}$ を状態遷移行列と呼ぶ. 観測値 x_t は現時刻の状態 z_t のみに依存して

$$p(x_t | z_t) = \prod_{k=1}^K p(x_t | \theta_k)^{z_{t,k}} \dots\dots\dots (8)$$

として生成される. ここで, θ_k は状態 k の確率モデルのパラメータである.

HMM を用いることで系列データのモデル化が可能となる. ところが, HMM では教師データ無しで系列データのクラスタリングを行うことはできない. 各系列データには系列データの属するクラスを示す正しいラベルが必要となる. そこで我々は前節で述べた GMM を利用して, 系列毎に GMM の各構成要素に属している確率を算出し, 最も確率の高かった構成要素のインデックスを教師ラベルとして HMM でのモデル化を行った. Table 1 に 129 個の系列デー

タの教師ラベルを示す.

それぞれのラベルに割り当てられた系列データを利用して HMM のモデル化を行う. HMM の学習には Hidden Markov Model (HMM) Toolbox for Matlab²¹⁾ を利用した. ラベル 4 については, 129 系列中割り当てられた系列データが一つだけであったため, 学習を行わなかった. HMM には状態数 5 の left-to-right HMM を採用し, その状態遷移行列を Baum-Welch アルゴリズムで学習した. HMM の出力はガウス分布に基づくとして, 初期値には GMM で用いた構成要素の平均値と分散を利用した. また, 状態遷移行列の初期値は上三角行列を仮定し, 対角成分とその次の列をそれぞれ 0.5 に設定した. ただし, 終了状態を示す状態 5 については自状態への遷移確率を 1 と設定した. 学習された遷移行列はラベル 1~3 についてそれぞれ,

$$\begin{pmatrix} 0.560 & 0 & 0 & 0 & 0 \\ 0.440 & 0.643 & 0 & 0 & 0 \\ 0 & 0.357 & 0.878 & 0 & 0 \\ 0 & 0 & 0.122 & 0.854 & 0 \\ 0 & 0 & 0 & 0.146 & 1 \end{pmatrix} \dots\dots\dots (9)$$

$$\begin{pmatrix} 0.975 & 0 & 0 & 0 & 0 \\ 0.025 & 0.975 & 0 & 0 & 0 \\ 0 & 0.025 & 0.979 & 0 & 0 \\ 0 & 0 & 0.021 & 0.724 & 0 \\ 0 & 0 & 0 & 0.276 & 1 \end{pmatrix} \dots\dots\dots (10)$$

$$\begin{pmatrix} 0.965 & 0 & 0 & 0 & 0 \\ 0.035 & 0.980 & 0 & 0 & 0 \\ 0 & 0.020 & 0.989 & 0 & 0 \\ 0 & 0 & 0.011 & 0.988 & 0 \\ 0 & 0 & 0 & 0.012 & 1 \end{pmatrix} \dots\dots\dots (11)$$

となった. ラベル 1 の系列を用いて学習した HMM 1 は, ラベル 2 の系列で学習した HMM 2 と比べて次状態への遷移確率が大きい. これは HMM 1 が比較的短い系列データの生成過程を良く表現していることを示し, 各 HMM はそ

Table 1 Number of sub-sequences in each class

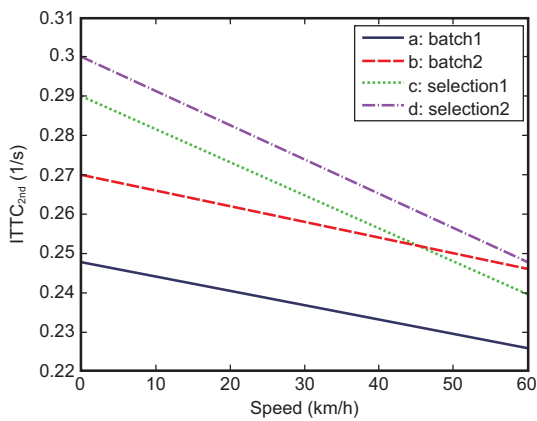
	トライアル数
ラベル 1	86
ラベル 2	26
ラベル 3	16
ラベル 4	1

それぞれ異なる運転状況に対応したブレーキ系列のモデル化を行うことができていると考えられる。

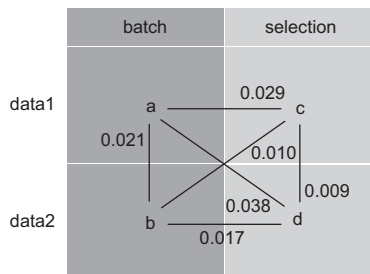
新しい系列データが観測されたとき、これらのHMMの尤度を算出することで、観測された系列データがどのクラスに属するかを判定することができる。

これらのHMMを評価するために、学習時とは別のデータセットを利用して判別率を調査した。上記手法による系列データのクラスタリングと、学習時と同様にGMMで作成した教師ラベルとの一致率は86.4%となった。更に、実際の走行時の映像を確認して各モデルがどのような運転状況に対応しているのかを確認した。その結果、ラベル1のブレーキは「接近」のブレーキ時系列、ラベル2は「信号待ち」、ラベル3は「右折」のブレーキ系列を示すことが確認できた。

このHMMを利用して、異常データのフィルタリングが可能となる。系列データに対して計算されるHMMの尤度はそのHMMから観測された系列データが生成される尤もらしさを表す。この尤度が一定のしきい値以下であった系列データに関しては通常の運転では生成されえないデータ、即ち異常データとして検出することが可能となる。



(a) Discriminating lines



(b) Comparison of the areas

Fig. 6 Differences of the areas of four OCSVMs

5. 数値実験

前節で説明したHMMを用いた異常データ判別の効果を確認するために、下記の数値実験を行う。同一条件で観測された三つのデータセットを準備する。ここでは、市街路で平日の午前中に計測された運転挙動データを利用した。三つの内、一つのデータセットを利用してHMMの学習を行い、残りの二つのデータセットを利用してこの評価を行う。

2節で述べたように、OCSVMでは95%のデータが存在する領域を分離するような判別超平面を学習する。ここでは二つのデータセットから四つのOCSVMを学習する。その内二つは訓練データを全て利用してOCSVMを学習したもの、残りの二つは先述したHMMによる異常データのフィルタリングを適用したものである。この四つのOCSVMの判別超平面の比較を行う。

Fig. 6 (a)に四つのデータセットから学習した判別直線を示す。ここで、HMMによるフィルタリングを行った場合の判別直線(c, d)は、そうでないもの(a, b)と比べて近い位置に存在している。これはHMMによる異常データのフィルタリングにより、OCSVMによる判別超平面の学習のロバスト性が向上していることを示す。Fig. 6 (b)に各判別直線間の距離を示しておく。

6. おわりに

ここでは、ドライバ適応型衝突警報システムのための異常データフィルタリング手法について述べた。ブレーキ系列の尤度に基づいて異常データのフィルタリングを行うこの手法により、警報システムにおいて警報を行うタイミングを決定する判別直線の学習を、データセットの違いに対してロバストに行うことが可能となった。ここで提案した手法と判別直線のオンライン学習を組み合わせ、さらに適応性を高めることは今後の課題である。

<参考文献>

- 1) 北島創, 安部原也, 伊藤誠, 平岡敏洋, 丸茂喜高, 和田隆広. 前方障害物に対する衝突リスク評価の検討. 自動車技術会学術講演前刷集, Vol. 106, pp. 13-16, 2008.
- 2) 味間弘喜, 池田和司, 柴田智弘, 深谷直樹, 人見謙太郎, 坂東誉司. TTC2ndとOCSVMを用いたドライバの運転状況判別. 自動車技術会, 秋季大会, 2009.
- 3) S. Kitajima, N. Kubo, T. Arai, T. Katayama. Reproduction of rear-end collision risk based on data acquired by drive video recorder and verification of driver's brake operation. JSAE Transaction, Vol. 39,

- pp. 205-210, 2008.
- 4) P. Barber and N. Clarke. Advance collision warning systems. IEE Colloquium, Vol. 234/2, pp. 1-9, 1998.
 - 5) J. Piao. Advanced Driver Assistant Systems from Autonomous to Cooperative Approach. Transport Reviews, Vol. 28, pp. 659-684, 2008.
 - 6) H. Mima, K. Ikeda, T. Shibata, N. Fukaya, K. Hitomi, T. Bando. Estimation of driving phase by modeling brake pressure signals. Proceedings of International Conference on Neural Information Processing, pp. 468-475, 2009.
 - 7) H. Mima, K. Ikeda, T. Shibata, N. Fukaya, K. Hitomi, and T. Bando. A Rear-end collision warning system for drivers with support vector machines. IEEE Workshop on Statistical Signal Processing, pp. 650-653, 2009.
 - 8) T. Kumagai and M. Akamatsu. Prediction of human driving behavior using dynamic Bayesian networks. IEICE Transaction on Information and Systems, pp. 857-860, 2006.
 - 9) J. C. MacCall and M. M. Trivedi. Driver behavior and situation aware brake assistance for intelligent vehicles. Proceedings of the IEEE, Vol. 95, No.2, 2007.
 - 10) V. N. Vapnik. The nature of statistical learning theory. Springer-Verlag, New York, NY, 1995.
 - 11) B. Schoelkopf, J. Platt, J. Shawe-Taylor, A. J. Smola, R. C. Williamson. Estimating the support of a high-dimensional distribution. Neural Computation, Vol. 13, pp. 1443-1471, 2001.
 - 12) C. C. Chang, C. J. Lin. LIBSVM – A Library for Support Vector Machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
 - 13) B. Swiderski, S. Osowski, A. Cichocki, A. Rysz. Single-class SVM and directed transfer function approach to the localization of the region containing epileptic focus. Neurocomputing, Vol. 72, pp. 1575-1583, 2009.
 - 14) S. Kitajima, Y. Marumo, T. Hiraoka, and M. Itoh. Comparison of evaluation indices for estimating driver's risk perception of rear-end collision. JARI Research Journal, Vol. 30, pp. 495-498, 2008.
 - 15) D. N. Lee. Theory of visual control of braking based on information about time-to-collision. Perception, Vol. 5, pp. 437-459, 1976.
 - 16) K. Ikeda, T. Aoishi. An asymptotic statistical analysis of support vector machines with soft margins. Neural Networks, Vol. 18, pp. 251-259, 2005.
 - 17) G. Fung and O. L. Mangasarian. Data selection for support vector machines classifiers. Proceedings of Knowledge Discovery and Data Mining, pp. 64-70, 2000.
 - 18) K. Igarashi, C. Miyajima, K. Itou, K. Takeda, F. Itakura, and H. Abut. Biometric identification using driving behavioral signals. IEEE International Conference on Multimedia and Expo, Vol. 1, 2004.
 - 19) T. K. Moon. The expectation-maximization algorithm. IEEE Signal processing magazine, Vol. 13, No.6, pp. 47-60, 1996.
 - 20) L. R. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings on the IEEE, Vol. 77, No.2, pp. 257-286, 1989.
 - 21) K. Murphy. Hidden markov model (HMM) toolbox for matlab. Online at <http://www.ai.mit.edu/~murphyk/Software/HMM/hmm.html>

<著 者>



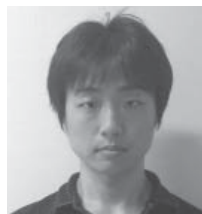
坂東 誉司
(ばんどう たかし)
研究開発3部
ドライバモデリング・データマイニ
ングに関する研究開発に従事



池田 和司
(いけだ かずし)
奈良先端科学技術大学院大学
情報科学研究科 教授
機械学習, 脳情報科学の研究に従事



味間 弘喜
(みま ひろき)
奈良先端科学技術大学院大学
情報科学研究科 博士前期課程修了
機械学習の研究に従事



井上 裕太
(いのうえ ゆうた)
奈良先端科学技術大学院大学
情報科学研究科 博士前期課程
在籍中
漫然運転警報システムの研究に従事



柴田 智弘
(しばた ともひろ)
奈良先端科学技術大学院大学
情報科学研究科 准教授
ロボットを用いた人の理解や支援に
関する融合領域研究に従事



人見 謙太郎
(ひとみ けんたろう)
株式会社トヨタ IT 開発センター
研究開発部
知能情報処理に関する技術調査・
研究に従事



深谷 直樹
(ふかや なおき)
研究開発3部
ドライバモデリング・データマイニ
ングに関する研究開発に従事