

## パターン認識による映像自己診断方式

榎本 暢 芳<sup>†,††</sup> 滝沢 圭<sup>†</sup> 谷口 恭 弘<sup>†††</sup>

一般環境下での映像が認識に適切か否かを診断する映像自己診断処理手法を提案する。対象とする映像は、異常有無の特定は可能であっても判別に有効な特徴を特定しにくいものであり、また診断処理には高速性が求められる。そこで提案手法では、あらかじめ異常有無のみを教示した映像サンプルに対して、非線型クラスタリングと判別分析学習を行っておき、未知入力に対しては判別空間上での k-最近傍識別 (k-NN 法) を実施する。また未知入力に対する汎化性を保つ目的でのクラスタ数決定に AIC を用いる。以上を国内道路での収集映像により評価した結果、照合精度 94%以上、かつ通常の k-最近傍識別法に比べ処理速度は約 100 倍となることを確認した。

### A video self-diagnosis method by using pattern recognition

NOBUYOSHI ENOMOTO<sup>†,††</sup>, KEI TAKIZAWA<sup>†</sup>  
and YASUHIRO TANIGUCHI<sup>†††</sup>

This paper presents a diagnosis method for video sequences in general condition. It is supposed hard to point out which factors in these video sequences are dominant to prevent object recognition. To solve this problem, a diagnosis system is learned by using non-linear clustering and discriminant analysis for sampled video sequences which show "Normal" or "Abnormal", and discriminates unknown inputs. To regularize our system to unknown inputs, we propose to decide number of clusters by AIC. By the test with video sequences captured on domestic roads, the proposed method classifies the sequences with over 94 % rate in 100 times faster than the normal k-NN method.

#### 1. はじめに

侵入者監視システム<sup>1), 2), 3)</sup> や車載映像監視システム<sup>4)</sup> の実現を目的として、一般のビデオ映像中から対象物を見つけて判定するための研究が行われている。このようなシステムへの入力映像は、環境変化の影響を受けるため、認識に不適格なものとなる場合があり、自己診断を行うことが必要と考えられる<sup>6)</sup>。またこのような映像は、パリエーションに富んでいるため、異常の有無を特定することは可能であっても、どの特徴量が判別に有効なのかを特定することが困難な場合があり、一般に線形識別が可能とも限らない。また自己診断は、本来の認識処理の前処理であるため高速動作が要求される。

今回対象とする映像のように複雑で非線型性のある入力を識別する手法として、サポートベクターマシン (SVM)<sup>5)</sup> が有名である。これは 2 クラス分類用線形識別器だが、Mercer カーネル<sup>7)</sup> と組み合わせると非線型問題の識別用に拡張できる。SVM ではその識別関数がつくる超平面を、サンプルのクラス境界の midpoint に配置するため、未知入力に対する汎化性が大きい。反面、一般に他の線形識別器に比べて処理時間がかかるという問題がある。

サンプルデータが多数そろえられ、かつ入力特徴次元数が大きくない場合には、k-最近傍識別法 (k-NN 法) は性能が高く有効な方法<sup>8)</sup> である。特に誤識率が Bayes 識別の最適誤り率の 2 倍以下になるという優れた性質がある<sup>9)</sup>。一方この手法には、入力特徴の次元数が大きかったりサンプルデータが膨大だった場合には、莫大な記憶容量と処理時間とを要するという問題がある。

以上の問題を解決するために、本論文ではサンプルデータからの特徴選択後に k-NN 法を用いる。また対象映像シーンについての正常、異常の境界は複雑な形状をしていることが予想されるため、まずサンプルベ

† (株) 東芝 社会ネットワークインフラ社  
Social Network & Infrastructure Systems Company,  
TOSHIBA CORPORATION

†† 電気通信大学  
The University of Electro-Communications

††† (株) 東芝 研究開発センター  
Corporate Research & Development Center, TOSHIBA  
CORPORATION

クトルを各クラス内において非線型クラスタリングした後に特徴抽出をおこなって、境界付近のデータ同士をなるべく離れさせるような写像を行う。さらに少数ベクトルによって元のサンプル分布を代表することによる汎化性低下を最小限に抑えるために、情報量規準を用いた学習制御方法を提案する。

以下2章において、上述したサンプル映像の学習による映像自己診断方法について説明し、3章においては国内の道路にて撮影した映像を用いた本方式の性能と学習制御方法の効果とを確認する。最後にまとめと今後の課題について4章にて述べる。

## 2. サンプル映像の学習による映像自己診断

### 2.1 入力空間分割と判別空間生成による次元圧縮

提案手法では、k-NN法の前段でサンプルデータを非線型クラスタリングしてサブクラスに分割した後、特徴選択を行うことで、低次元の代表ベクトルで表現する。

特徴選択は、入力サンプル vector を  $X$ 、ある線形変換行列 (特徴選択行列) を  $A$  とすると以下で示される。

$$Y = AX \quad (1)$$

特徴選択行列  $A$  は、サンプルに対して、これが上記サブクラスに分割されているとの仮定のもとで、線型判別分析 (LDA) を実施して算出する。

識別段階においては、入力ベクトル  $X$  から式1によって選択された特徴ベクトル  $Y$  に対して、各サブクラスの代表ベクトル  $R_c$  ( $c$  はサブクラス、ここでは  $R_c$  はサブクラスの重心) との間でユークリッド距離により k-NN 識別を行う。

今もし複数のベクトルがサンプルデータとして与えられており、それぞれが”正常”、”異常”のいずれかのクラスに分類されていたとする。またこれら2クラスが前述のクラスタリング結果としてそれぞれ  $C_n$  個、および  $C_a$  個、合計  $C$  個のサブクラスに分割されていたとする。この時、式1の各サブクラスに対する特徴選択行列  $A$  は、クラス間分散  $B$ 、およびクラス内分散  $W$  を用いた下式の一般化固有値問題の解として求められ、各サブクラスに対して級間分散の級内分散に対する比率を最大化する。すなわちある入力について、この行列を用いると、各サブクラス同士を引き離す方向に変換する。

$$(B - \lambda W)A = 0 \quad (2)$$

ここに変換後のベクトル次元数  $d$  の最大値は  $C - 1$  である。

上記を踏まえて識別時における演算時間を積和回数で見積もると下表となる。ここで  $N$ : サンプルベクトル数、 $n$ : 入力  $X$  の次元数、 $m$ : 特徴選択ベクトル次元数である。

$k - NN$	$nN$
$LDA + k - NN(2 \text{ クラス})$	$nm + 2N$
$Clustering + LDA + NN$	$nm + mC$

表1 様々な特徴選択と組み合わせた k-NN の積和回数

一般に  $m \ll N$  であるから表中段の2クラスのみでの識別が高速だが、特徴選択ベクトルの次元数が1となり性能が著しく低下する。一方、提案手法では、代表ベクトル次元数  $d$  が小さくなり過ぎない範囲でクラス数  $C$  を選択することで性能劣化をおさえながら、 $nm + mC \ll nN$  と処理速度向上が期待できる。

### 2.2 非線型クラスタリング

入力ベクトルを非線型クラスタリングする方法としては実装がシンプルな自己組織化写像 (SOM)<sup>11)</sup> を用いる。

SOM はニューラルネットによるクラスタリング手法で、その構造は1次元、2次元などの低次元に配列したノードの結合によって表され、各ノード  $n_i$  にはノードを代表する重みベクトル  $W_i$  を持つ。そして SOM の処理では入力ベクトル  $X$  と最も類似する重み  $W_i$  をもつノード (勝者ノード) を選択し、その重みを入力ベクトルに次第に近づけるようくり返し更新する。

## 3. 情報量規準による学習パラメータ最適化

入力クラスタリングでの、分割クラス数は学習の汎化性能と関係があり、クラスタが多いほどサンプルデータを正確に表現できるが、クラスタが多過ぎる場合にはサンプルに適合しすぎて汎化能力が低下する。ここではクラスタ数の選択を行うため、情報量規準の一つである AIC を用いる。AIC はモデルの最大経験対数尤度を  $\log L$ 、最尤推定するモデルの自由度を  $M$  とすると以下で表され、その最小 (極小) 値をとる場合において、学習データに対する当てはまりの良さと、自由度の小ささ (汎化性の高さ) とがバランスする。

$$AIC = -2 \log L + 2M \quad (3)$$

ここでは SOM によるクラスタリングにおける指定クラス数を様々に変えた場合に生成される各サブクラス領域がそれぞれ正規分布に従うと仮定する (類似する研究に k-means アルゴリズムのクラス数決定

がある<sup>12)</sup>。そして全領域についての AIC を極小とする場合を最適なクラス分類として選択する。各サブクラス領域  $r_i$  内のデータが  $n$  次元正規分布で表されており、サブクラス数が  $C$  とすると全領域の AIC は以下となる。

$$AIC = -2\sum_{i=1}^C L(\mu_i, \Sigma_i) + 2C(n + \frac{n(n+1)}{2}) \quad (4)$$

ここに  $\mu_i$ 、 $\Sigma_i$  は領域  $i$  の平均ベクトル、分散共分散行列であり、 $L(\mu_i, \Sigma_i)$  は対数尤度である。

#### 4. 性能評価

##### 4.1 対象とした映像サンプル

性能評価に用いた対象映像シーンは 2000 年 11 月に都内周辺道路において車両走行中に 8mm VTR にて撮影した約 2 時間分の映像であり、全画像は 108522 フレーム、うち正常、異常と教示された画像はそれぞれ 60825 フレーム、47710 フレームである。

図 1 に映像シーン中の画像例を示す。

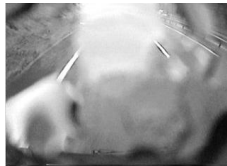


図 1 異常な映像の一例 (雨滴)

学習用サンプルは、前述の正常、異常と教示されたものそれぞれについて、ランダムにその 1/10 を選択して利用したが、識別用特徴量 (ベクトル) としては、画素の輝度頻度分布、画素間差分の輝度頻度分布、連続 2 フレーム間での対応画素の差分輝度頻度分布、および画面内特定領域周辺の画素平均値とし、その次元数は 99 次元とした。

サンプルベクトル群を主成分分析 (PCA) して、第一主成分と第二主成分とに投影したものを図 2 に示す。緑色領域は”正常”クラス、赤色領域は”異常”クラスであり、両クラスの境界は複雑な形状をしていることがわかる。

##### 4.2 サンプルベクトルの非線型クラスタリング

前々節で述べたサンプルベクトル群を”正常”クラス、”異常クラス”のそれぞれについて 2 次元 SOM を用いて事前クラス数を  $2 \times 2 = 4 \sim 5 \times 5 = 25$  とし、くり返し回数を 10000 としてクラスタリングした結果を以下図 3、図 4 に示す。分割クラスを増やす

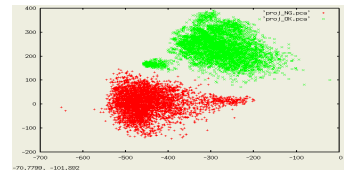


図 2 学習サンプルの PCA 表示

に従って、PCA 平面上で境界領域に存在したベクトルが境界付近でなかったベクトルと非線型に融合してクラスを作ってあり、2 クラスの境界を遠ざける働きををすると思われる。

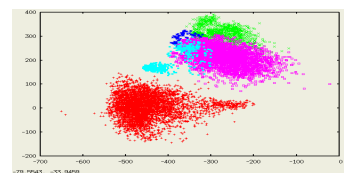


図 3 学習サンプルのクラスタリング (2 × 2)

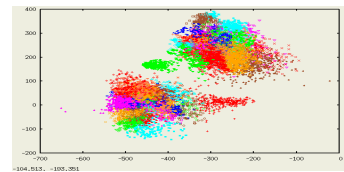


図 4 学習サンプルのクラスタリング (5 × 5)

#### 4.3 識別結果

前節の全映像について、サンプルを除いた画像について識別評価を行った。

対比用の  $k-NN$  法 (単純  $k-NN$  法) による識別結果を表 2 に、本論文で提案する非線型クラスタリングと判別分析による識別結果を表 3 にそれぞれ示す。ここで提案手法による判別空間での  $k-NN$  識別では簡単のため  $k = 1$  とした。

k の値	誤識別率	フレーム平均処理速度
1	0.68%	0.33sec
10	1.91%	0.32sec
100	5.90%	0.32sec

表 2 k を変えたときの k-NN 法の性能

表 2 から単純  $k-NN$  法では  $k$  が大なるほど誤識別率が上昇することが、表 3 から提案手法では分割数が大なるほど誤識別率が低下することがわかる。

また、”正常”/”異常” 両クラスをそれぞれ 64 個に分割した場合の提案手法の識別性能が単純  $k-NN$  法

正常クラス数	異常クラス数	誤識別率	平均処理速度
4	1	29.2%	0.00008sec
9	9	17.2%	0.00018sec
16	16	11.9%	0.00031sec
25	25	8.91%	0.00059sec
36	36	7.81%	0.00144sec
49	49	6.47%	0.00256sec
64	64	5.78%	0.00382sec
81	81	5.52%	0.00382sec
100	100	5.35%	0.00839sec

表 3 分割数を変えたときの提案手法の性能

において精度の低い  $k = 100$  とした場合に相当しており、単純  $k-NN$  法での識別性能の高さが確認できる。一方で各フレームの平均処理速度を比較してみると、提案手法では最も時間のかかる分割数  $100 \times 100$  の場合以外では、単純  $k-NN$  法に比べて処理速度が 100 倍程度となっている。

次に以下図 5 では両クラスを上述表 3 の各分割数でクラスタリングした場合のそれぞれについて AIC を算出したものと誤識別率との推移を示す。青色線は”正常”クラスについて、緑色線は”異常”クラスについての AIC の推移であり、赤色線は誤識別率の推移である。図中分割数が大きくなると誤識別率が急激に減少するが、分割数 = 64 付近を越えるあたりからあまり低下しなくなる。また”異常”クラスについての AIC は分割数 = 9 の周辺で極小となっているが、これは誤識別率の低下が最も急激な分割数と符合する。一方、”正常”クラスについての AIC は分割数 = 64 の周辺で極小となっているが、これは誤識別率の低下が一定値に収束しだしている分割数と符合する。これは提案手法においては、サンプルの分割クラスタに対する AIC を用いることにより、未知入力に対する誤識別率を制御可能であることを意味する。この理由は、今回提案手法のモデルが小サンプルによって大量の未知入力を識別するものなので、モデル汎化性が誤識別率を直接反映しているためと思われる。

## 5. おわりに

本論文では、一般環境下での認識処理に対する入力映像の診断を行う手法を、非線型クラスタリングと線形判別分析の組み合わせとして提案した。道路収集映像を用いた評価から、照合精度 94% 以上、処理速度は通常の  $k-NN$  法の 100 倍程度となることを確認した。また、この実験からサンプル分割クラスタに対する AIC の極小値を用いて誤識別率の制御が可能となることを確認した。

今後は本応用に対して他の非線型識別手法を適用し

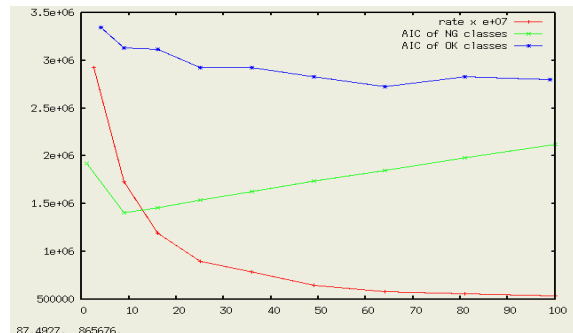


図 5 分割数による AIC と誤識別率の推移

た場合との、より詳細な比較が必要と思われる。

## 参考文献

- 1) Lipton, Fujiyoshi and Patil, "Moving Target Classification and Tracking from Real-time Video" IEEE Workshop on Applications of Computer Vision (WACV), Princeton NJ, October 1998, pp.8-14.
- 2) Collins, et al "A System for Video Surveillance and Monitoring: VSAM Final Report", Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May, 2000.
- 3) N. Enomoto, et al, "A method for monitoring activities of multiple objects by using stochastic model," Proc. MVA2000 IAPR Workshop on Machine Vision applications, Nov 2000.
- 4) Taniguchi, Okamoto, "Automatic Rear and Side Surveillance System using Image Processing," 6th World Congress on Intelligent Transport Systems, 1999.
- 5) Vapnik, V., "Statistical Learning Theory," Wiley, 1998.
- 6) N. Enomoto, K. Takizawa, "Novel functions for Automatic Rear and Side Surveillance," Proc. IV2001.
- 7) Müller, et al, "An introduction to kernel-based learning algorithms," IEEE Transactions on Neural Networks, 12(2), 181-201, 2001.
- 8) 麻生, 津田, 村田, "パターン認識と学習の統計学," 岩波書店, 2003.
- 9) Cover, T.M. and Thomas, J.A., "Elements of Information Theory," Wiley Interscience, 1991
- 10) Akaike, H., "A new look at the statistical model identification," IEEE Trans. on Automatic Control, vol.AC-19, No.6, pp.716-723, 1974
- 11) T. Kohonen, "Self-Organization and Associative Memory," Springer-Verlag, 1992.
- 12) Hardy, A., "On the Number of Clusters," Computational Statistics & Data Analysis 23, pp.100-108, 1996