

## クラスタリングと Boosting を用いた高速、高精度な映像自己診断方式の提案

榎 本 暢 芳<sup>†,††</sup> 富 田 悦 次<sup>††</sup> 仲 川 崇 史<sup>††</sup>

本論文では、一般環境下での認識処理に対して診断要因とその程度とを陽に指定せずに入力映像の診断を行うことが可能な映像自己診断処理手法を提案する。映像監視のような一般環境化での認識用入力映像の診断は、診断対象映像の特徴がバ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容易であっても、具体的にどの特徴量によって判別すべきかを特定することは困難である。このような一般の複雑な入力に対する判別性能の点で k-最近傍識別法がすぐれているが、必要とする記憶容量と計算量が多い。そこで正常、異常各カテゴリに属するサンプル映像について事前にクラスタリングと判別分析学習を行って精度劣化をおさながらカテゴリを代表するパラメータ数を減少させる。さらに各カテゴリ中のクラスタ間距離に基づいてカテゴリを識別する弱識別器を複数構成し、それらの boosting 学習を行うことでさらに精度向上をはかった。国内道路にて撮影した映像を用いて評価した結果、判別精度 97%以上、かつ処理速度は同一学習データを利用する k-最近傍識別法の 50 倍以上となることを確認した。また一般的な非線型サポートベクターマシンによる判別性能と比較して、判別精度、処理速度ともに上回ることを確認した。

### A video self-diagnosis method by using an optimized classifier with boosted clusters of learning patterns

NOBUYOSHI ENOMOTO,<sup>†,††</sup> ETSUJI TOMITA<sup>††</sup>  
and TAKASHI NAKAGAWA<sup>††</sup>

This paper presents a diagnosis method for video sequences in general condition, for which we does not need to specify important image features explicitly. It is supposed hard to point out which factors in these video sequences are dominant to prevent object recognition. Instead of this, we can say whether detection method works for each input image by simulation using sampled images. To solve these problem, a diagnosis system is learned by sampled video sequences which show "Normal" or "Abnormal" in advance, and discriminates unknown inputs. The k-nearest-neighbours-classification rule(k-NN) is known as a good method for these problems. Meanwhile this method needs huge memories and computation time. The proposed method uses k-NN on a discriminant space which is made by clustering and linear discriminant analysis for classifying video sequences with small memory in real time. Moreover, we propose to use boosting for making combined learner from weak learners which are composed from the clusters. By the test with video sequences captured on domestic roads, the proposed method classify the sequences with over 97 % rate in the faster speed than those of k-NNs, and a kernel based SVM.

#### 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、屋外での撮影映像を中心とした一般のビデオ映像の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、時々刻々収集環境が変化しながら収集されるために、認識に適さないものが入力される場合

がある。したがって、本来の認識処理に先だって入力映像に対する自己診断を行うことが必要と考えられる<sup>5)</sup>。ところで、このような一般環境化での認識用入力映像の診断は、入力映像の画質が自動認識に適するか否かという 2 つの状態を判別する問題と考えられるが、診断対象映像の特徴がバ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、状況に応じてどの特徴量が判別に有効かを特定することが困難な場合が少なくない。また本来の認識処理の前処理として適用するものであるため、高速な識別処理が要求される。さらにこれら入力は複雑で、識別困難な場合も少なくない。

非線型性を含むような複雑な入力に対する識別機と

<sup>†</sup> (株) 東芝 産業システム社  
Industrial Systems Company, TOSHIBA CORPORATION  
<sup>††</sup> 電気通信大学  
The University of Electro-Communications

して、近年サポートベクターマシン (SVM)<sup>5)</sup> が注目を集めている。これは 2 クラスの分類問題を解くための線形識別器であるが、サンプルを Mercer カーネル<sup>7)</sup> によってより高次元空間に非線型写像を行っておき、この空間内での学習を行うことによって、非線型問題の識別を行うように拡張できるものである。線形 SVM の識別関数は以下で表される。

$$f(x) = \sum_{j=1}^d w_j x_j + b \quad (1)$$

ここに  $w_j$  は重みベクトル、 $b$  はバイアス項であるが、入力ベクトルが  $d$  次元とした場合には  $f(x) = 0$  は  $d - 1$  次元の超平面を形成する。SVM ではこの識別関数のパラメータ  $w_j$  と  $b$  に対し、この超平面が分割する 2 クラス境界領域に存在する各クラスのサンプルの内、最も近いもの同士の中点でバランスするように学習を行う。これによってサンプルが少数しか無い場合においても、未知入力に対する汎化性を最大化している。このように優れた SVM であるが、特に非線型 SVM は、処理に多くの時間を必要とするという問題がある。

一方サンプルデータが多数そろえられ、かつ入力特徴の次元数があまり大きくない場合には、 $k$ -最近傍識別法 ( $k$ -NN 法) の識別性能が高いことが知られている。これはサンプルベクトルを全て記憶しておき、入力とサンプルベクトルとの距離の近いものから  $k$  個を選択して、それらの多数決によって入力の存在するクラスを推定する<sup>8)</sup> ものである。例えば  $1 - NN$  法の識別誤り率  $R_{1NN}$  は、以下のように Bayes 識別の最適誤り率  $R_B$  の 2 倍以下であるという優れた性質がある<sup>9)</sup>。

$$R_B \leq R_{1NN} \leq \left(2 - \frac{K}{K-1}\right) R_B \quad (2)$$

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必要とするという問題がある。

この問題を解決するために、本論文ではサンプル入力データ (サンプル入力ベクトル) から特徴選択を行うのと同時に、複数収集されたこれらサンプル入力ベクトルを代表する少数のデータ (代表ベクトル) を用いて各カテゴリを表現し、それらを使った判別を行うことが有効と考える。

また一般に診断対象となる映像シーンを正常カテゴリと異常カテゴリとに判別する場合に、その境界領域は複雑な形状をしていることが予想されるため、特に境界付近に存在する学習サンプルデータもそれ以外のサンプルデータと同様に正しいカテゴリに判別されるように判別用代表ベクトルを選択することが望ましい

と考えられる。

このような機能を実現するために、前処理としてサンプルベクトルを各カテゴリ内においてクラスタリングし、その結果を線形判別分析することで各クラス代表ベクトルを算出しておく。さらに各カテゴリから選択したクラス代表ベクトルとの距離を用いる識別器の組を複数構成し、その出力加重を用いて総合的にカテゴリを判別することを考える。

このとき学習段階において識別器の選択と加重和とを適切に求めておくことは、判別処理時間を減少させるとともに、判別精度向上の意味で重要である。本論文では、このような操作を AdaBoost<sup>10)</sup> による組み合わせ最適化の問題に帰着させて解決することを提案する。

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て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

本論文で対象とする一般環境で収集された映像の自己診断を行う問題の解決策として、 $k$ -NN 法のように入力とカテゴリ中のサンプルデータとの距離を直接評価するのではなく、あらかじめ入力、および各サンプルそれぞれから特徴選択を行うのと同時に、サンプルデータを代表する少数、かつより低次元の代表ベクトルを算出して評価し、その結果を用いてカテゴリを判別する方法が判別時の処理速度の面で有効と考える。また今回対象とする映像監視用データには、各カテゴリ中に広範囲に分布するデータを含む可能性があるため、なるべく認識に有効な特徴を選択して利用する方が判別性能の面からも有効であると考える。

このような特徴選択の中で処理が単純なものは線形変換であり、入力サンプル vector を  $\mathbf{X}$ 、ある線形変換行列 (特徴選択行列) を  $\mathbf{A}$  とすると

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \quad (3)$$

のように行われる。

特徴選択行列  $\mathbf{A}$  は、学習段階において各カテゴリ中の入力サンプルをあらかじめ複数クラスに分割 (クラスタリング) しておき、それらをクラスとして扱った線形判別分析 ( $LDA$ ) を実施することで算出しておく。入力ベクトルをクラスタリングする方法としては種々考えられるが、ここでは実装がシンプルなクラスタリング手法として、データマイニングやデータ可視化に広く利用されている  $K$ -平均法 ( $K$ -means clustering method) を用いる。

識別段階においては、入力ベクトル  $\mathbf{X}$  から式 (3) によって選択された特徴ベクトル  $\mathbf{Y}$  に対して、各クラスの代表ベクトル  $\mathbf{r}_c$  ( $c$  はクラス、ここでは  $\mathbf{r}_c$  はクラスの重心) との間でユークリッド距離  $D_c = |\mathbf{Y} - \mathbf{r}_c|$  を計算し、距離小なるクラス代表ベクトルが属するカテゴリに、入力ベクトルもまた属すると識別することにする。

このために”正常”, または”異常” の各カテゴリから各々一個のクラス代表ベクトルを組にして選択し、それぞれの代表ベクトルと特徴ベクトルとの距離を比較して識別特徴量とする。このとき各組が構成する識別器を  $f_j(\mathbf{X})$  とするとき、識別特徴量を以下の 2 値で出力することとしておく。

$$f_j = 1 \text{ if } D_n < D_a \quad (4)$$

$$f_j = -1 \text{ if } D_a < D_n \quad (5)$$

ただし  $D_n$ : 正常クラス代表ベクトルと特徴ベクトル  $\mathbf{Y}$  との距離,  $D_a$ : 異常クラス代表ベクトルと特徴ベクトル  $\mathbf{Y}$  との距離である。

そして最終的な判別は上記各識別器ペアの出力特徴量の加重和として以下を評価することで求める。

$$F(\mathbf{X}) = \text{sign}\left(\sum_{j=1}^C (W_j f_j(\mathbf{X}))\right) \quad (6)$$

ここに  $\text{sign}(a)$  は  $a$  の符号を表す関数とし,  $C$  は後述のように正負両カテゴリのサンプルデータをクラスタリングした時のクラスタ総数である。

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, 少なくともそのそれぞれが”正常”, または”異常” のいずれかのカテゴリに分類されていたとする。またさらにこれら 2 カテゴリ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  $C_a$  個のクラスに分割されていたとする。この時, 各クラスに対するクラス間分散  $B$ , およびクラス内分散  $W$  は以下で示される。

$$B = \frac{1}{N} \sum_{c=1}^C N_c (\bar{x}_c - \bar{x})(\bar{x}_c - \bar{x}) \quad (7)$$

$$W = \frac{1}{N} \sum_{c=1}^C \sum_{i=1}^{N_c} (x_{ic} - \bar{x}_c)(x_{ic} - \bar{x}_c) \quad (8)$$

$$C = C_n + C_a \quad (9)$$

ここに  $N$  はサンプル総数,  $C$  は”正常”, ”異常” カテゴリのそれぞれから生成されたクラスの総数,  $N_c$  は各クラス内のサンプル数,  $\bar{x}_c$  はクラス  $c$  内のサンプル平均ベクトル,  $\bar{x}$  はサンプル全体の平均ベクトル, そして  $x_{ic}$  はクラス  $c$  内の第  $i$  ベクトルである。

上記  $B, W$  を用いると, 式 (3) で示した特徴選択行列  $A$  は, LDA における特徴選択行列となり, 下式に対する一般化固有値問題の解として求められる。そし

てこの行列は, 各クラスに対して級間分散の級内分散に対する比率 (判別規準) を最大化する。すなわちある入力について, この行列を適用した結果は, 各クラス同士をより引き離す方向に変換するものとなる。

$$(B - \lambda W)A = 0 \quad (10)$$

クラスの代表ベクトル  $\mathbf{r}_c$  は, 前述のとおり各クラス  $c$  の重心ベクトル  $\mathbf{g}_c$  ( $c = 1, 2, \dots, C$ ) を式 (3) によって変換したものとするが, 変換後のベクトル次元数  $d$  は以下のように両クラスの総和より 1 だけ小さくなる。

$$d \leq (C_n + C_a) - 1 = C - 1 \quad (11)$$

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。サンプルベクトル数を  $N$ , 入力  $\mathbf{X}$  が  $n$  次元, 特徴選択ベクトル  $\mathbf{Y}$  が  $m$  次元 ( $m < n$ ) とすると  $k$ -NN 識別 ( $k = NN$ ), および 2 カテゴリを  $C$  個にクラスタ分割後, 判別空間上で  $k$ -NN 識別する方法 (*Clustering + LDA + k - NN*), および今回提案方法 (*proposed*) の処理速度は下表となる。

表 1 様々な判別方法の積和回数

$k - NN$	$nN$
<i>Clustering + LDA + k - NN</i>	$nm + mC$
<i>proposed</i>	$nm + m((C_n C_a))$

ここで  $nN$  は各サンプルベクトルとの距離計算のために,  $nm$  は特徴選択のために, さらに  $mC$  は各サブクラスの参照ベクトル  $\mathbf{r}_c$  との距離計算のために必要となる。一般に  $m \ll N$  であるが, 代表ベクトル次元数  $d$  を小さくし過ぎない範囲で  $C \ll N$  のようにクラス数を選んだ場合には  $nN \ll nm + mC$  となるため, 今回提案の方法と判別空間上で  $k$ -NN 識別する方法とにおいて, 性能劣化をおさえながら処理速度の向上が期待でき, 処理速度向上のオーダは両者のうち判別空間上で  $k$ -NN 識別する方法が上回ることが予想される。

## 2.2 boosting による学習パラメータ最適化

今回提案の方法では, 式 (6) に示すように正常, および異常クラスの代表ベクトルを組として構成した識別器の出力特徴の加重和としている。

前節では特徴選択によるデータ量圧縮について述べたが, これら識別器の組み合わせ選択と重み付けとを学習サンプルに基づいて最適に決定することでこのようなデータ圧縮による性能劣化を挽回することを考える。

具体的には上述の各識別器の出力特徴から各々弱識別器を構成し, この結果を boosting の中でも識別性能の高い Real AdaBoost を用いて統合し, 式 (6) に相当する判別式を求めることとした。以下にその学習

手順を示す。

(初期化)

for  $i = 1$  to  $N$  do  $D(i) = 1/m$  end

ただし  $D(i)$ : 入力  $i$  の確率分布,  $N$ : 全サンプルベクトル数とする。

(学習)

for  $t = 1$  to  $T$  do

(1) 全サンプル入力に対し弱識別器応答を決定

(i) 正解サンプル入力での特徴量頻度分布生成

$$\begin{aligned} F(y_i = 1|f_j) &= \sum_{(i=1, \dots, N | \mathbf{x}_i \in f_j \wedge y_i = +1)} (D(i)) \\ &= W_{j,+1} \end{aligned}$$

ここに  $f_j = -1, 1$ : 特徴量,

$y_i$ : 入力の真値 (1: 正常, -1: 異常),

$F(y_i = 1|f_j)$ : 頻度分布。

(ii) 異常サンプル入力での特徴量頻度分布生成

$$\begin{aligned} F(y_i = -1|f_j) &= \sum_{(i=1, \dots, N | \mathbf{x}_i \in f_j \wedge y_i = -1)} (D(i)) \\ &= W_{j,-1} \end{aligned}$$

ここに  $F(y_i = -1|f_j)$ : 頻度分布。

(iii) 損失が最小になる弱識別器を選択

以下で損失  $Z$  を最小とするように、

全識別特徴の組みについて

弱識別器  $h_t(\mathbf{X})$  を選択する。

$$\begin{aligned} h_t(\mathbf{X}) &= \frac{1}{2} \log \left( \frac{F(y_i = 1|f_j)}{F(y_i = -1|f_j)} \right) \\ Z &= 2 \sum_{j=1}^{C_n} C_a \sqrt{F(y_i = 1|f_j) F(y_i = -1|f_j)} \end{aligned}$$

(2)  $h_t(\mathbf{x})$  を使って、入力確率分布  $D_t(i)$  を更新

$$D_{t+1}(i) = D_t(i) \exp(-y_i * h_t(\mathbf{X}_i))$$

end

(強識別器生成)

最終的な強識別器  $H(x)$  を生成し、今回提案の判別の使用する。

$$H(x) = \text{sign} \left( \sum_i (h_t \mathbf{X}) - b \right)$$

ただし  $b$ : 閾値である。

### 3. 性能評価

本章では、前章にて述べられた映像自己診断処理手法について、国内道路にて撮影した映像を用いて性能の評価を行った結果を示す。

#### 3.1 対象とした映像

対象とした映像シーンは 2000 年 11 月に東京都府

中市、および調布市周辺の一般道と高速道路とにおいて車両走行中に後方視界を 8mm アナログ VTR にて撮影した約 2 時間分の映像である。またそのデータは 8bit/pixel の gray 階調画像としてデジタル化されたもので、そのサンプリング間隔は 15fps(frames per second) である。したがって全画像枚数は 108522 フレームであり、このうち目視により正常と教示された画像は 60825 フレーム、異常と教示された画像は 47710 フレーム存在した。

図に映像シーンのうちのいくつかの画像を示す。



図 1 正常な映像の一例



図 2 異常な映像の一例 (雨滴)



図 3 異常な映像の一例 (逆光)

#### 3.2 映像学習用サンプルベクトル

映像の学習に用いたサンプル用の画像としては、前述の 108522 フレームのうち、正常と教示されたもの、異常と教示されたものそれぞれについて、全体の 1/10 の枚数を選択して利用した。したがって正常、異常それぞれについて 6083 フレーム、4771 フレーム、合わせて 10854 フレームの画像を利用することになる。

実験では、以上の映像シーン中の各画像から特徴量として単純で一般の映像に共通な以下を抽出し、学習用サンプルベクトルとして用いた。これらは、画面内画素の輝度頻度分布、画面内の画素間差分の輝度頻度分布、時間的に連続する 2 画面内での対応画素の差分輝度頻度分布とした。また車載応用を考慮したため、各輝度頻度分布は走行車線の左右両側の領域について算出した。さらに応用時の高速応答性を考慮して、通常画素階調数に相当する 256 次元である頻度分布は各々 16 次元に間引いて使用することにした。以上から

サンプルの各ベクトルの次元数は  $16 \times 3 \times 2 = 96$  次元となる。

### 3.3 単純 1-NN 法による識別結果

本節では、本論文で提案する自己診断手法との対比を行うために、前節で述べたサンプルを用いた単純 1-NN 法による識別結果を示す。

表 2 k を変えたときの k-NN 法の性能

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

ここにこの手法において各カテゴリの代表となるベクトルは前説で述べた学習用サンプルフレームから抽出されたベクトルを、そのまま用いている。

また評価用入力ベクトルは、前節にてサンプルとして抽出したフレームの残りのフレームをさらに 1/10 に分割した 9766 フレームからなる 10 シーンのうちの 1 シーンから抽出したものをを用いた。特に近傍数  $k = 1$ 、つまり 1-NN 識別を行った場合などは、性能としては誤識別率 0.68% と良好であるが、処理速度は 0.3sec/frame を要している。今回対象として監視系は、入力として 15fps(=0.06sec/frame) 程度の映像を対象としているので、単純 1-NN 法では、約 5frame ごとにしか映像監視制御を実施できないことになるため、このような自己診断には向かないことがわかる。

### 3.4 他手法と提案手法での識別結果比較

以下には前章において提案手法に対して、次元圧縮によって処理速度を向上可能な類似手法である *Clustering + LDA + k-NN*、識別率面で優れた手法の代表である SVM、および今回提案する *Clustering + LDA + boosting* について誤識別率、各フレームを処理するのに要する平均時間を評価した結果が表 3 である。ここに今回提案手法として、線形判別分析基底数 (LDA 基底数) を 96 としたものと、寄与率上位から 29 のみを利用したもの 2 通りを評価した。

ここに各手法の学習用サンプル入力ベクトルとしては、前節で用いたものを全て同様に用いた。

一方、評価用入力ベクトルは、前節と同様に抽出した 9766 フレームからなるシーン 10 系列から抽出したものをを用い、識別率、フレーム平均処理時間は各系列ごとに得られた結果の平均値を求めた。

なお、各手法で用いている識別用のパラメータは以下であるが、SVM については線形のもの、RBF カーネルを用いるものの 2 種類を評価対象とした。そのパラメータとして前者についてはスラック変数が、後者としては RBF 係数  $\gamma$  があるが、それぞれ入力サンプルについて最適パラメータを実験的に求めたものを用いた。

さらに SVM には様々な実装が存在するが、今回評

表 3 LDA+k-NN, SVM, および提案手法の性能比較

手法	誤識別率	フレーム平均処理時間
Clustering +LDA +k-NN (ただし k=1)	3.74%	0.0040sec
線形 SVM	6.5%	2.05E-6sec
非線型 SVM	3.1%	0.0048sec
Clustering +LDA +Real AdaBoost (LDA 基底数=96)	2.44%	0.0065sec
Clustering +LDA +Real AdaBoost (LDA 基底数=29)	2.36%	0.0038sec

価には  $SVM^{light11}$  を使用した。

表 4 LDA+k-NN, SVM, および提案手法のパラメータ

手法	パラメータ	パラメータ値
Clustering +LDA +k-NN (ただし k=1)	クラスター数	正常, 異常 カテゴリとも各 50 個
線形 SVM	スラック変数	0.5
非線型 SVM	スラック変数	0.17
	RBF パラメータ	0.000165
Clustering +LDA +Real AdaBoost (LDA 基底数=96)	クラスター数	正常, 異常 カテゴリとも各 50 個
	学習くり返し数 t	5200
Clustering +LDA +Real AdaBoost (LDA 基底数=29)	クラスター数	正常, 異常 カテゴリとも各 50 個
	学習くり返し数 t	6842

以上の結果から今回提案手法は 2 例ともに、識別率の面で処理時間を要する  $k-NN$  法に比べてやや不足するが実用範囲内であり、処理速度は約 50 倍高速であった。

*Clustering + LDA + k-NN* との比較では、提案手法のうち基底を 96 次元用いたものでは識別性能では上回ったが、処理速度ではわずかに及ばなかった。ただし前章で予想したほどの処理速度の差とはなっていない。これは本手法が使用する弱識別器数は最大では正常クラス数と異常クラス数との積となるが、実際はより少ない識別器数で学習が収束するためである。さらに判別分析の段階で基底数を圧縮して 29 次元とした場合は、識別性能、処理速度ともに *Clustering + LDA + k-NN* を上回った。

識別性能で定評のある SVM と比べた場合、識別性能ではいずれの提案手法でも線形、非線型 RBF カーネルの両手法より上回ることが確認できた。処理速度の面では線形版 SVM よりかなり (1000 倍以上) 低速

であるが、提案手法のうち 96 基底を用いたものでも RBF カーネル版よりはわずかに及ばない程度であり、29 基底にしたものでは RBF カーネル版 SVM を上回ることを確認した。ここから LDA を用いて元もとの特徴量を圧縮した後でも、boosting による特徴選択の効果が発揮され、性能を挽回できていることがわかる。

一方で今回提案手法についての学習くり返し数  $t$  を確認してみるといずれの基底数を用いた場合でも、収束までに 5000 回を越える学習を必要としており、学習時間を要する手法であると考えられる。

#### 4. おわりに

本論文では、一般環境下での認識処理に対して診断要因とその程度とを陽に指定せずに入力映像の診断を行うことが可能な映像自己診断処理手法を提案し、その処理手順について、クラスタリングと線形判別分析に boosting を組み合わせて定式化し、識別性能を劣化させずに処理速度を向上することを目的とした。提案手法については国内の道路収集映像を用いて評価を行い、自己診断における判別精度 97% 以上、かつ処理速度は同一学習データを利用する  $k$ -最近傍識別法の 50 倍以上となることを確認した。さらに識別性能において定評のある SVM との比較を実施し、一般的な RBF カーネルを用いた非線型 SVM に対して、提案手法が識別率、処理速度の両面で上回る性能を達成できることを確認した。今後は別応用への適用可能性をも検討することが必要と考えられる。

また前節でも述べたように、本手法は今回の応用については学習時間を要することが確認されているが、学習時間の対象依存性解析、学習高速化が課題である。

#### 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), pp.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) Y. Freund, and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, 55(1), pp.119-139, 1997.
- 11) T. Joachims, "Learning to Classify Text using Support Vector Machines," Kluwer Academic Publishers, 2002.