Boost 学習に基づく特徴量の貢献度を用いた特徴選択手法

土屋　成光†　　藤吉　弘亘†

† 中部大学大学院工学研究科 〒 487-8501 愛知県春日井市松本町 1200
E-mail: †{tsuchiya,hf}@vision.cs.chubu.ac.jp

あらまし　画像認識分野では AdaBoost，Support Vector Machines(SVM) 等の識別器が用いられている．これらの識別器は入力特徴量が性能に大きく影響するが，認識に対する特徴量の有効性評価は難しい．そのため，多数の特徴量に対する有効性の自動評価は有用である．そのような問題に対して，SVM 識別器のマージンを用いて入力特徴セットを評価し，特徴選択を行うことが提案されている．しかし，マージンの値は計算上，完全に分離ができる入力よりも不完全な入力の方が大きな値となり得る．そこで，本稿では Boost 学習を用いた特徴量評価法と，それを用いた特徴選択法を提案する．まず SVM の識別性能と貢献度の相関性を調査し，有効性を確認した．次に，特徴選択を行い，Confident Margin(CM) による特徴選択法との比較を行った．その結果より，CM では正確に評価できないケースに対して提案手法が正しく評価でき，提案する特徴選択法が有効であることを示した．

キーワード　特徴選択，特徴評価，ブースティング，物体識別，機械学習

A Method of Feature Selection Using Contribution Ratio based on Boosting

Masamitsu TSUCHIYA† and Hironobu FUJIYOSHI†

† Dept. of Computer Science, Chubu Univ. Matsumoto 1200, Kasugai, Aichi, 487-8501 Japan
E-mail: †{tsuchiya,hf}@vision.cs.chubu.ac.jp

Abstract　AdaBoost and Support Vector Machines (SVM) are popular in the field of object recognization. Classification performance of these classifiers is sensitive to the feature set. In order to improve classification performance, in addition to choosing between the classification, we have to pay attention to which subset of features to employ in the classifier. To solve this problem, a method of evaluating feature set using a margin of decision boundary from SVM classifier has been proposed. However, the margin of SVM sometimes will be large caused by outliers. In this paper, we propose a method of feature selection using contribution ratio based on boosting. We show that the contribution ratio is effective for evaluating features. Comparing to the conventional method by confident margin, the proposed method can select better feature set using the contribution ratio obtained by boosting.

Key words　feature selection, feature evaluation, boosting, object recognization, machine learning

1. はじめに

近年，画像認識分野では一般に，物体やシーンを認識するために，画像から特徴量を計算し，それを入力とした識別器を構築する方法が採られている．識別器には，非線形な識別境界が構築可能なものが用いられている．特に，識別時の高速性や，弱識別器の変更による問題への柔軟性から Boosting 手法 [1] を用いるものが増加している．Viola らは，空間と時間変化に着目し，Haar-like wavelets を用いた AdaBoost による人の検出法を提案している [2]．また，Hoiem らは Colour，Texture，Location，3D-Geometory に大別される特徴を入力として Boosted decision trees を構築し，画像中の三次元構造を復元する手法を提案している [3]．また，同様の特徴量を用いた車両，人認識が提案されている [4]．

人や自動車の認識等の問題では，同一対象における見えの違い，車種等の種別の多様性に影響を受けない特徴量を選択する必要から，入力特徴の選定は非常に重要な問題である．

特徴量を評価，選択する方法としては，実際に識別器に特徴を追加，削除した際の識別率の比較，全サンプルの分散をより大きく持つ特徴軸を抽出する主成分分析（principal component analysis, PCA），クラス間の分離をより容易にする特徴軸を抽出する重判別分析（multiple discriminant analysis, MDA）な

どが知られている．しかし，実際に入力特徴を増減しての識別率の比較は，識別器構築の際に選択する学習用データと評価用データの偏りに大きく影響を受けることや，特徴量の数だけ識別器の構築，識別実験を行う必要があるという問題がある．また，PCA は全サンプルを表現する低次元の近似空間を生成する手法であるため，その評価値は非線形な識別器を構築する際の有効性を直接示すものではない．MDA は線形分離可能な対象に対しては，クラス間の識別に有用な評価が可能である．しかし，対象の問題が線形，非線形どちらであるか，また，識別境界が何次関数か不明な場合，安易に評価することができない．

一方，青木らは非線形な識別境界を構築可能である Support Vector Macine(SVM) のマージンに，誤識別率より得られる信頼度（Confidence）の値を考慮して得られる Confident Margin を用いて特徴選択を行うことを提案している [9]．これは，ある入力特徴セットで構築された SVM による識別器がどの程度良好な性能を持つかを評価する，いわば非線形識別器の解析により得られる特徴量の評価，選択手法である．

そこで，本稿では Boost 学習により実際に得られる弱識別器とその重みを用いて，簡易に非線形な識別境界の構築に際して有用な特徴評価を行うことを提案する．Boosting による識別器は，そのアルゴリズムによって有用な特徴量を多く，不要な特徴量を少なく識別に利用するよう構築される．これは，Boosting による識別器自身による特徴評価であるといえる．データの分散による評価方法と異なり，対象の問題について詳細に調べることなく簡易に特徴評価を行うことができる．また，本稿では，得られた貢献度の確からしさを示すため，2 種類の評価実験を行う．まず人工データ，実問題それぞれに対して貢献度を算出し，SVM により各特徴量を削減した識別器の識別性能と比較を行い，特徴量の貢献度が識別性能に相関性を持つこと，また，他の手法で構築された識別器にも応用可能であることを示す．次に，Confident Margin による特徴選択法との比較として，2 種の問題に対して特徴選択実験を行い，有効性を示す．本稿の内容は，6 つの章で構成されている．まず 1 章は，全体のまえがきである．2 章で従来法について述べ，従来法の特徴を記述する．3 章で，特徴選択時に指標とする Boost 学習に基づく特徴評価法について述べ，4 章は，Boost 学習に基づく特徴選択法について述べる．5 章は，従来法として挙げる Confident Margin による特徴選択法との比較実験である．最後の章は，全体のむすびである．

2. 従来の特徴選択法

本章では，識別器への入力特徴を評価する従来の手法について述べ，その特徴と問題点について述べる．

2.1 データの分散を用いる入力特徴の評価法

データの分散を考慮する特徴量評価法として主成分分析（Principal Component Analysis），重判別分析（Multiple Discriminant Analysis）を比較に用いる．

2.1.1 主成分分析：PCA

主成分分析とは，図 2 に示すように入力特徴量で構成された空間を全てのデータの分散を最大とするように直交な主成分空間で近似する方法である．主成分分析によって，元の情報を保ちながら次元数を削減することができる．また，部分空間に射影することにより，射影された特徴量は，全体の特徴ベクトルを自動的に重み付けされたものとなる．この際，生成される主成分はそれぞれ特徴空間の情報量をどれだけ保持しているかを示す寄与率という値を持ち，主成分の持つ累積寄与率と，主成分と特徴量の相関係数である負荷量の大きさによって特徴量を評価可能である．

2.1.2 重判別分析：MDA

判別分析とは図 2 のように，クラス間の分散を最大にし，クラスの分離が容易な成分を求めるものである．判別分析を多群（multi class）に対応させたものを重判別分析（multiple discriminant analysis）という．判別分析では一般に，Wilks の Λ 統計量より得られる偏 F 値を用いて入力特徴を評価する．

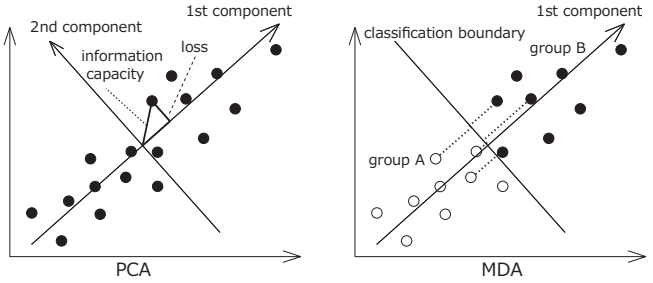

図 1　従来の特徴評価法

これらデータの分散に基づく手法では，どの特徴量が非線形識別器の入力として実際にどの程度有効であるかは不明である．

2.2 識別器の解析に基づく特徴選択法

識別器の解析に基づく特徴選択手法として，Confident Margin による SVM のための特徴選択が青木らにより提案されている [9]．マージンの値が大きい，すなわち各クラスの端点と識別境界が離れている識別器は，より汎化性の高い分離性能を持つことが期待できる．そのため，マージンを用いて SVM 識別器の識別性能や入力特徴の評価が可能となる．しかし，マージンの値は計算上，線形分離が不可能な場合にも非常に増大することが報告されており [9]，これに対し青木らは誤識別に基づくペナルティを与えることによって誤選択を抑止できるとしている．しかし分離が困難などのケースによっては分離性能の低い特徴量を用いた際のマージンが分離性能の高い特徴を用いた際のマージンの数 10 倍等と極端に大きくなり，抑制が困難な可能性がある．これは，マージンを評価に用いることの計算上の危険性と困難さを示唆しており，実際にこの値を用いる際は，対象がそのような評価困難な問題であるかどうかを調査する必要がある．これらの評価値は識別器の性能を表わすものであるが，識別性能の値に過敏であり，実際にどの特徴量がどれだけ有効なのかを把握し難いという問題がある．

3. Boost 学習に基づく特徴量の貢献度評価法

特徴選択は内部の特徴量や識別器に対してなんら知見なく行えることが望ましい．その点において PCA, MDA ではデータの持つ非線形性や複雑性についての考慮が必要であり，SVM

のマージンより評価を与えることは誤選択の危険性がある．また，マージンによる評価は特徴セットに対して評価を与える手法であるため，識別率を求める必要がないといっても 1 特徴量の削減のために特徴量数と同数の識別器の構築が必要となり，Confident Margin においては Confidence の算出のために識別関数の出力が全データ分必要となる．これは，特徴量数が非常に増大している昨今を顧みると，近い将来に非効率となる可能性がある．そこで我々は，識別性能に依存しない特徴選択のための指標として Boost 学習による識別器に着目する．Boost 学習による識別器は，そのアルゴリズムによって有用な特徴量を多く，不要な特徴量を少なく判断に利用するよう構築される．これは，Boost 学習による識別器自身による特徴評価である．Boost 学習を行う際の弱識別器を各特徴量と関連付け，学習の結果得られる重みを考慮することで，実際に識別に対してその特徴量が貢献した度合いを求める．この値は，距離概念等が介在しない非常にシンプルな値である．しかしながら，特徴量を用いて実際に対象に対する識別器を構築するために必要な度合が具体的に情報として表わされる．これにより，一度の識別器構築でその段階での全特徴を評価し，不要な特徴量を見つけることが出来る．弱識別器は 1 特徴量のみを用いた識別器とする．

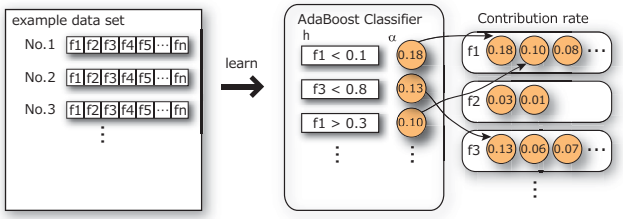

図 2　Boost 学習に基づく特徴量の貢献度評価

3.1　AdaBoost による学習

各識別対象毎に特徴量の評価を行うため，単純に 2 クラス識別器 $H(\boldsymbol{x})$ を $H^A, H^B, H^C, H^D \cdots$ のように各クラス $class = \{A, B, C, D, \cdots\}$ に対して構築し，各クラス毎に学習過程で選択された弱識別器に着目する．まず各クラスに対して，対象クラスかそうでないかを出力とする 2 クラス識別器 H^c を構築する．c は対象クラスのラベルである．このとき，他クラスの positive サンプルを自クラスにおいて negative サンプルとして学習する．

AdaBoost の学習による弱識別器 h の生成は，まず学習サンプルから両クラスのヒストグラムを特徴量 p 毎に求め，次に，それら全てに対して誤識別率を最小とする閾値 th を探索することで行う．学習回数 t 回目における弱識別器 $h_t^{p,th}(\boldsymbol{x})$ は，図 3 中（c）のように $t-1$ 回目の学習結果に基づいて重み付けされたデータに対して，特徴量集合 P 全てについて閾値探索を行い，最良の識別率が得られる特徴量 p とその閾値 th を採用する（式 1）．

$$h_t^{p,th}(\boldsymbol{x}) = \underset{(0 \lessgtr th,\ p)}{argmin} \{\sum_i \delta_K(y_i,\ \mathrm{sgn}\ (x_i^p - th))\} \qquad (1)$$

AdaBoost では，識別関数 $h_t(\boldsymbol{x})$ が誤識別を起こしたデータを重視して再学習を行う．この処理をラウンド数 T 回反復し，生成された識別器群の識別関数のアンサンブルによって最終的な識別関数を生成する．最終的な識別関数 $H_T(\boldsymbol{x})$ は次式で表される．

$$H_T(\boldsymbol{x}) = \sum_{t=1}^{T} \alpha_t h_t(\boldsymbol{x}). \qquad (2)$$

ここで，α_t とは t 番目の弱識別器による識別結果の信頼度を表し，最終的な識別に対して弱識別器 h_t の結果が影響する度合いである．以上のアルゴリズムを図 3 に示す．

Algorithm The Adaboost algorithm

1. **Input:** n, Training dataset $(\mathbf{x}_i, y_i)$
2. **Initialize:** $w_1(i) = 1/n (i = 1 \cdots n)$, $h_0(\mathbf{x}) = 0$
3. **Do for** $t = 1, \ldots, T$. $\epsilon_t(h) = \sum_i^n I(y_i \neq h(\boldsymbol{x}_i)) w_t(i)$
   (a) $\epsilon_t(h_{(t)}) = min \epsilon_t(h)$
   (b) $\alpha_t = \frac{1}{2} ln(\frac{1-\epsilon_t(h_{(t)})}{\epsilon_t(h_{(t)})})$
   (c) $w_{t+1}(i) = w_t(i) \exp(-\alpha_t h_{(t)}(\boldsymbol{x}_i) y_i))$
4. **Output:** Final hypothesis with weights α_t

$$sign(H_t(\mathbf{x})),\ where\ H_T(\boldsymbol{x}) = \sum_{t=1}^{T} \alpha_t h_t(\boldsymbol{x})$$

図 3　Adaboost 学習アルゴリズム

3.2　特徴量の貢献度

弱識別器として選択された特徴量と，その弱識別器の正規化後の重みである α' は，その識別対象クラスの識別に重要な要素である．そこで，学習後の AdaBoost において選択された各特徴量の弱識別器から，貢献度 CR_p を以下のように定義する．

$$CR_p = \sum_{t-1}^{T} \alpha'_t \cdot \delta_K[F(h_t), f] \qquad (3)$$

ここで，α' は識別器において重み α を次式のようにその総和で正規化したものとする．

$$\alpha'_t = \frac{\alpha_t}{\sum_{i=0}^{T} \alpha_i} \qquad (4)$$

$F(h_t)$ は，h_t に採用された特徴量を求める関数である．この CR_p を各クラス識別器において選択された弱識別器群から求めると，特徴量 f_p がその識別器の識別能力にどの程度貢献するかを示す指標となる．このような特徴量の貢献度を予め求めることで，より多くの学習サンプルを用いて他の識別器（SVM 等）を構築する際の特徴選択の指針とできる．

Algorithm The feature contribution evaluating algorithm

1. **Input:** *feature set* $\boldsymbol{f} = [f_1, f_2, ..., f_n]$,
   trained AdaBoost classifier $H_T(\boldsymbol{x}) = \sum_{t=1}^{T} \alpha_t h_t(\boldsymbol{x})$
2. **Initialize:** $\boldsymbol{CR} = [CR_1, CR_2, ...CR_n] = 0$
3. **Do for** $p = 1, \ldots, n$.
   $CR_p = \sum_{t-1}^{T} \alpha'_t \cdot \delta_K[F(h_t), f]$
4. **Output:** Feature contribution rate $\boldsymbol{CR}$,
   Contribution rate of feature p : CR_p

図 4　貢献度評価アルゴリズム

3.3 貢献度の評価実験

従来法との比較として，特徴量に対する従来法，提案手法による各評価値と，その特徴量を除いた際の SVM 識別器の識別性能の変化の相関性によって評価を行う．相関性が高いほど有効な評価値であるといえる．従来法としては，principal component analysis(PCA)，multiple discriminant analysis(MDA) を用いる．

PCA は，得られた寄与度を各特徴量のスコア C_p^{pca} へ式 5 によって変換する．因子負荷量 $L_{i,j}$ は第 i 特徴ベクトルと第 j 主成分ベクトルとの相関係数であり，N は成分数である．

MDA の評価値は特徴軸の判別への重要度を表す偏 F 値を用いる．

$$C_p^{pca} = \sum_{j=1}^{N} \lambda_j L_{p,j} \tag{5}$$

3.3.1 識 別 問 題

特徴評価を行う識別問題として 2 種の問題を扱う．

問題 1. 線形分離不可能なデータの識別

線形分離不可能な問題に対して提案手法が特徴量を評価可能であることを示すため，それぞれがオーバーラップを持つ，f1，f2，f3 の 3 種の特徴量を用いて構成した 2 クラス識別問題を用いる．データは各クラス 400 点を生成した．

問題 2. CU database による物体識別

実問題として，図 5 に示すような物体画像を自動車（VH）／人（SH）／複数の人（HG）／二輪車（BK）の 4 クラスに識別する問題である．識別器への入力特徴量としては，表 1 に示すような，領域全体より得られる大域的な特徴量 7 種類を用いる．

図 5 識 別 対 象

表 1 入力特徴セット：問題 2 CU database

特徴		ベクタ数
形状	縦横比と主軸の傾き (AS)	2
	複雑度 (CS)	1
テクスチャ	垂直方向エッジ (V)	2
	水平方向エッジ (H)	2
	右上がり方向エッジ (R)	2
	左上がり方向エッジ (L)	2
時間	オプティカルフローの分散 (OF)	1

3.3.2 比較評価結果

各問題に対する提案手法による評価値，PCA で得られた評価値，MDA で得られた評価値，識別性能の変化をそれぞれ特徴量ごとに図 6 に示す．図より，問題 1 において提案手法が最も識別性能の変化をよく表現できていることがわかる．また，問題 2 において従来法と提案手法はよく似た傾向を持つ．しかし，表 2 に示す相関係数の値をみると，提案手法が最も識別性能の変化を表現できていることがわかる．ここから，提案手法は従来法に対し表現力の高い，有効な特徴評価であると考えられる．

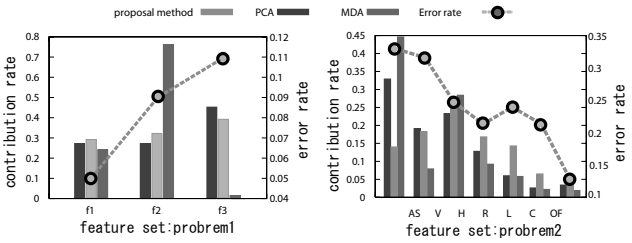

図 6 従来法との比較

表 2 各問題に対する評価値と識別性能の相関係数

	問題 1	問題 2	平均
提案手法	0.91	0.79	0.85
MDA	-0.11	0.63	0.26
PCA	0.76	0.55	0.655

4. 貢献度に基づく特徴選択

貢献度は入力特徴の有効性を評価したものである．図に示すように特徴量の貢献度に基づいて入力特徴の有効性を評価し，特徴選択を行う．

4.1 特徴選択アルゴリズム

特徴選択は貢献度評価の値が最も低かった 1 特徴量を段階的に取り除くことで行う．1 特徴量毎の削除という意味では同一であるが，Confident Margin による特徴選択法が特徴数と同数の特徴サブセットにおける Confident Margin 評価を行うことで 1 特徴を削除するのに対し，提案手法では 1 度の貢献度評価により 1 つの特徴量を削除する．これは，Confident Margin が特徴セットにより構築される識別器の性能評価値であるのに対し，貢献度が特徴セット内における各特徴量の有効性の評価値であるという，評価対象の相違に起因する．

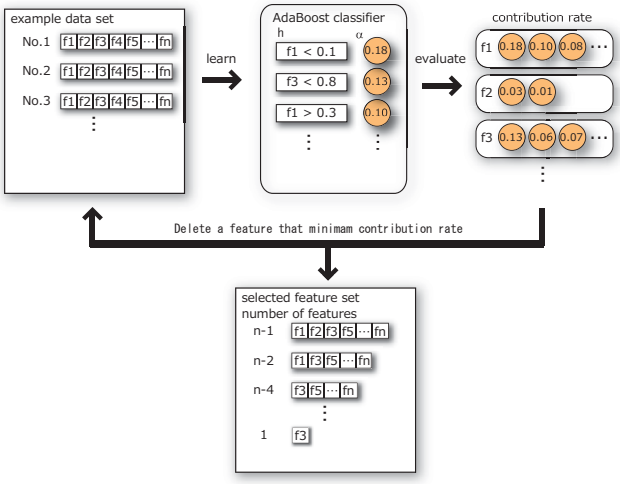

図 7 貢献度に基づく特徴選択

Algorithm The feature selection algorithm

1. **Input:** n, Training dataset $(\mathbf{x}_i, y_i)$
2. **Initialize:** Subset of surviving features $\boldsymbol{s} = [1, 2, ..., n]$
3. **Do for** Until $\boldsymbol{s}$ is empty
   (a) Train AdaBoost classifier
   with all the training exsamples
   (b) Compute Contribution Rate $\boldsymbol{CR}_1, \boldsymbol{CR}_2, ..., \boldsymbol{CR}_n$
   (c) find the worst feature
   $worst = argmin(\boldsymbol{CR}_i)$
   (d) Remove the worst feature i that minimam $\boldsymbol{CR}_i$

図 8　特徴選択アルゴリズム

4.2　比較評価実験

比較評価実験として，UCI データベース [11] 内の Sonar データ，CU データベース内の VH クラスの判別に対して Confident Margin，貢献度それぞれにより特徴選択を行う．Sonar データは 60 種類の特徴量，208 サンプルを持つ 2 クラス分類問題，VH データは 7 種類の特徴量，200 のポジティブサンプル，600 のネガティブサンプルを持つ 2 クラス分類問題である識別器は SVM 識別器を用い，実装は SVM-light により行った．学習の際のパラメータはすべての実験において，$c = 10$，$\sigma = 1.8$ とし，文献 [9] と同値を用いた．また，Confident Margin では学習が早期に飽和すると誤識別率の算出が困難なため，誤識別率は Leave-one-out によるクロスバリデーションにより算出している．

4.2.1　Confidence による抑制が十分なケース：Sonar データ

表 3 に，それぞれの最大性能（Best RR）とその際の特徴次元数（DIM）を示す．

表 3　Sonar データに対する特徴選択結果

	Best RR[%]	DIM
Proposal method	91	36
Confident Margin	93	13

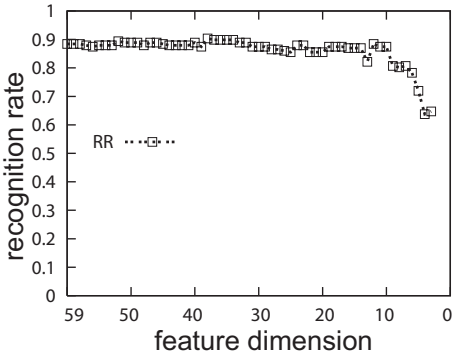

図 9　sonar データに対する実験結果

表より，提案手法では Confident Margin に比べ倍の特徴量を使った際に最大の識別性能が得られており，識別性能自体も 2%低い結果である．しかし，Sonar データに対する $\boldsymbol{CR}$ による特徴選択結果（図 9）より，提案手法においても特徴量数の削減を行いつつ，識別性能（Recognition Rate:RR）の大きな低下が少ないことがわかる．すなわち，有用な特徴の選別を行い，識別性能を保持したままの次元削減に成功している．

4.2.2　Confedence による抑制が十分でないケース：VH データ

提案手法による VH データに対する特徴選択結果を図 10（CR）に，Confident Margin による特徴選択による選択結果を同図中 RR（CM）に，また，その際の Confident Margin を同図中（CM）に示す．図より，両手法ともに特徴量数の削減を行いつつ，識別性能の大きな低下が少ないことがわかる．しかし，Confident Margin による特徴選択における識別性能は，最後に残った 1 特徴量のみで識別を行った際大きく低下し，提案手法により残った最後の 1 特徴量を用いて識別を行った際の識別性能に対して 2%以上の低下がみられる．同図中に示してある，Confident Margin（CM）の値に着目すると，最後の 1 特徴量として OF を選択した際，異常なほどに大きく上昇していることが見て取れる．この際の Confidence，マージンをそれぞれ表 4 に示す．表より，これは文献 [9] で述べられているように，誤識別を起こすデータが多いとマージンが非常に大きくなるというケースであると考えられる．実際，表からは誤識別に応じて算出される Confident の値は非常に小さく，マージン値を抑制する方向であることがわかる．しかし，マージンが 50 倍程度の差を持つためそれが吸収できず，有用でない特徴量 OF が最後まで残る結果となった．それに対して提案手法は，Confident Margin で最後まで残った特徴量を最初に除外し，識別率は保たれている．これは，実際に識別に用いられる特徴軸を評価する貢献度の値が，SVM のマージンのように誤識別データの特徴空間上の位置等によって大きく左右されない性質を持つためである．

表 5 に，それぞれの最大性能とその際の特徴数を示す．図 10，表 5 より，特徴選択課程において提案手法と Confident Margin に有意な差は見られない．しかし，先ほどの結果より，Confident Margin では有用でない特徴量が残り，提案手法では実際に識別に有効な特徴量しか残らなかったことがわかる．

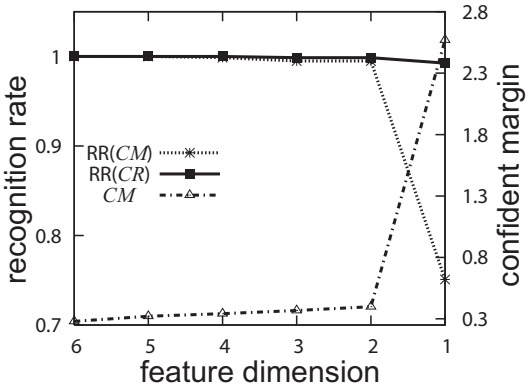

図 10　VH データに対する実験結果

表 4　DIM=1 の際のマージン

NM	c	CM
5.13	0.50	2.57

表 5　VH データに対する特徴選択結果

	Best RR[%]	DIM
Proposal method	100	3
Confident Margin	100	4

5. おわりに

Boost 学習により貢献度という新たな特徴量の評価値を得る手法を提案した．識別器の識別性能の変化との比較により，提案手法が PCA，MDA といった従来法より線形判別不可能，大域的な特徴量を用いたケースにおいて高い表現力を持つことを示した．また，貢献度を用いた特徴選択を提案し，比較実験を行った．その結果，Confident Margin による特徴選択に比べ識別性能の向上には大きく寄与できないが，Confident Margin では誤選択が発生するようなケースにおいても正しく選択が可能であることを示した．以上より，貢献度は物体識別のための特徴評価として有効な手法であり，それを用いた特徴選択は SVM マージンによらないため誤選択の少ない選択手法であるといえる．今後は，SFS(Sequential Forward Selection) [12]，SBS(Sequential Backward Selection) [13] に代表される一般的な特徴選択法 [14], [15] と比較し，パフォーマンスについて評価を行っていく予定である．また，SFS，SBS 等における貢献度の評価値としての利用について研究を進め，特徴間の共起性を表現した Boost 識別器，複数の Boost 学習手法間での評価，選択結果の比較にを行う予定である．

文　献

[1] 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, Aug, 1, 55, 119-139, 1997.

[2] P. Viola, M. J. Jones, D. Snow : “Detecting Pedestrians Using Patterns of Motion and Appearance”, Proc. of the Ninth IEEE International Conference on Computer Vision (ICCV’03) - Volume 2, 2003.

[3] D. Hoiem, A. A. Efros, M. Hebert: “Automatic photo pop-up”. ACM SIGGRAPH2005, pp.577-584, 2005.

[4] D. Hoiem, A. A. Efros, M. Hebert: “Putting Objects in Perspective”, CVPR2006, pp.2137-2144, 2006.

[5] M. Tsuchiya, H. Fujiyoshi : “Evaluating Feature Importance for Object Classification in Visual Surveillance”, Proc. of ICPR 2006; 18th Int’l Conf. on Pattern Recognition (HongKong/China), pp. 978–981, 4 pages (Aug. 2006).

[6] 小村剛史，藤吉弘亘，矢入郁子，香山健太郎，吉水宏：“歩行者 ITS のためのフレーム間差分による移動体検出法とその評価」，情報処理学会論文誌 CVIM, Vol.45, No.SIG13(CVIM10), pp.11-20, Dec. 2004.

[7] L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone, “Classification and Regression Trees.”, Wadsworth, CA, 1984.

[8] Jerome Friedman, Trevor Hastie and Robert Tibshirani. Additive logistic regression: a statistical view of boosting. The Annals of Statistics, 38(2):337-374, April, 2000

[9] 青木一真, 黒柳奨, マウリシオ クグレ, アント サトリヨ ヌグロホ, 岩田彰：「Confident Margin を用いたＳＶＭのための特徴選択手法」, 電子情報通信学会論文誌　Vol.J88-D-Ⅱ No.12 情報・システム Ⅱ パターン処理 , 2005.

[10] Nello Cristianini (原著), John Shawe - Taylor (原著), 大北 剛 (翻訳): ”サポートベクターマシン入門”，共立出版，2005.

[11] C. Blake, E. Keogh and C.J. Merz: UCI Repository of machine learning databases [http://www.ics.uci.edu/mlearn/MLRepository.html]. Irvine, CA: University of California, Department of Information and Computer Science, 1998.

[12] Whitney,A.W : ”A direct method of nonparametric measurement selection” IEEE Trans.Comput.20, pp.1100-1103, 1997

[13] Marill,T.and D.M.Green : ”On the effectiveness of receptors in recognition system” IEEE Trans.Inform.Theory 9, pp.11-17, 1963 from a single image. In Proc. ICCV, 2005.

[14] P.Pudil et al. : ”Floating Search Methods in Feature Selection” Pattern Recognition Letters, Vol.15, No.11, pp.279-283, 1994

[15] T. Bäck, U. Hammel, and H.-P. Schwefel, “Evolutionary Computation: Comments on the History and Current State,” IEEE Trans. Evolutionary Computation, vol. 1, no. 1, pp. 3-17, 1997