Boostingに基づく分割統治的戦略による高精度な識別器構築手法

土屋　成光[†a)]　　藤吉　弘亘[†]

A Method for Constructing a High-Accuracy Classifier Using Divide and Conquer Strategy Based on Boosting

Masamitsu TSUCHIYA[†a)] and Hironobu FUJIYOSHI[†]

あらまし　顔や人等の物体検出では，同一カテゴリーに所属しながらも多くの姿勢や見えの異なるインスタンスに対して検出を行う必要がある．これらカテゴリー内での変化は画像上で大きな影響をもたらし，異なる姿勢の物体全てを同一に学習することは困難となる場合がある．この問題を解決するために本論文では，学習サンプルをサブカテゴリーに分割し，それぞれに対して学習を行う分割統治的戦略に着目する．弱識別器応答の類似度に着目した学習サンプルのクラスタリング法と，それにより生成されたサブカテゴリー群に対する学習法としてJoint Boostingを改良したDivide-and-Conquer Boostingを提案し，高精度な検出を実現する．評価実験の結果，2クラスBoostingに比べINRIA Person Datasetを用いた評価では誤識別率1.0%において8.0%，UIUC Image Database for Car Detectionを用いた評価では誤識別率0.1%において9.5%検出率を向上することができた．

キーワード　物体検出，Boosting，Joint Boosting，分割統治法

1. まえがき

近年の物体検出は，画像局所特徴量と統計的学習手法の組合せにより実現され，統計的学習法による識別器にはSupport Vector Machines (SVM) やBoosting [1] などの利用が提案されている．特に，局所特徴量とBoostingの組合せは，高次元の入力から必要な特徴量を自動選択するため，有用な特徴量を解析しやすく，また特徴量削減効果もあるため，高次元の入力特徴を用意した人，顔，車両等の物体検出に多く用いられている [2]～[4]．検出対象となる物体のアピアランスは，照明，視点によって大きく異なり，例えば人や顔，車両などは向きや姿勢により，画像としては形状，テクスチャなど全く異なるものが得られる．それに対してDalalらは，Histograms of Oriented Gradients (HOG) 特徴量とSVMによる人検出を提案している [15]．HOG特徴量は，輝度こう配情報に基づき形状の人らしさを抽出する特徴であり，体型，服装等の個体差による問題を大きく低減することに寄与している．山内らは単純なHOGだけでなく，特徴間の共起表現とBoostingを応用してJoint特徴量という強力な特徴量を提案している [2]．高木らも同様に，Geometric Contextに共起表現を導入し，Boostingによって選択させることで高精度な人，車の検出を実現している [3]．しかし，人のような様々な変動を含む物体を検出する際，変化の大きい学習サンプルを一括に扱い識別器を構築するのは困難となる場合がある．文献 [2], [3] の手法は特徴量の表現力を向上させることで精度が向上しているが，学習自体はReal AdaBoostなどの既存の2クラス学習を用いており，その点について検討されていない．

図1は人検出用データベースとして一般的なINRIA person dataset [13] に含まれている画像である．全て人を中心に捉えるよう画像上で正規化されていながらも，体幹の傾き，顕著な年齢差，複雑な姿勢等変化に富んでいる．これら全てに共通した識別を行うのは非常に困難である．しかし，いくつかの集合に対してなら共通した傾向を見出すことができる．例えば，図1 (a)，(b) は全体的な傾き，頭，肩と背景の境界部分等に類似点がある．我々は，カテゴリー中の部分集

† 中部大学工学部情報工学科，春日井市
Department of Computer Science, Chubu University, Kasugai-shi, 487–8501 Japan
a) E-mail: tsuchiya@vision.cs.chubu.ac.jp

 © （社）電子情報通信学会 2012

図 1　カテゴリー内での変化 (INRIA person dataset)
Fig. 1　Variation in a category (INRIA person dataset).

合であるサブカテゴリーに着目する．

全学習データから通常の 2 クラス識別器を学習するだけでなく，サブカテゴリーそれぞれに対して個別に働く識別器を学習することで，複雑な問題を効率的に解くことが可能である．本論文では学習サンプルの偏りをサブカテゴリーとみなして自動分割（図 3 (a)）し，それらを小問題として個別に問題を解く識別器を導入（図 3 (b)），その解を統合することで主問題を効率的に解く分割統治的アプローチを提案する．従来，このような複雑な変化をする検出対象に対する解決策としては Viola らの顔検出 [4] に倣うことが多い．この文献は Haar-like 特徴量と Boosting を用いた高速，高精度な顔認識手法を提案しており，顔の向きなどに対する頑健性を得るため顔の傾き，方向ごとに顔画像を人手により分類し，サブカテゴリーごとに別個に識別器を構築する方法を述べている．また，Huang らは Viola らと同様に顔画像を手動で分類し，多クラス識別器に入力することで同様に頑健性を担保している [5]．しかし，これら人手による分類はサブカテゴリーを付与するコストが非常に大きい上に，その性能はサブカテゴリーの設定が的確であるかどうかにより大きく左右される．例えば図 1 (h) はやや正面を向いていながらも全体としては左向きである．このような，どのラベルを付与すべきか判定が困難な例が存在する．また，図 1 (f) は正面を向いているため頭，肩等において図 1 (a)，(b) と同じ特徴といえる．しかし，体幹が傾いているため脚部が左下方向に伸びている点は左向きの画像である図 1 (c) と同じ特徴ともいえる．サブカテゴリーごとに分割して識別器を構築する場合，各識別器は「サブカテゴリー（右向き，左向き等）らしさ」を学習する中で知識として，「対象カテゴリー（顔，人など）らしさ」について同様の情報をそれぞれ重複して学習していると考えられる．これら重複した情報の学習は計算コストを引き上げる原因となる．

そこで我々は検出対象となる学習サンプルを自動的にサブカテゴリーに分割し，分割されたサンプルに対してサブカテゴリー固有の情報は固有に，カテゴリー全体やカテゴリー群に共通する情報は共有して識別器を構築する分割統治のアプローチを提案する．既に，複数のカテゴリーを対象とした問題に対する共有的な多クラスの Boosting として，Joint Boosting が提案されている [7]．Joint Boosting は，高精度な 2 クラス識別器である Gentle Boost [6] を多クラスに拡張したものであり，一般物体認識の分野において統合的な多クラス識別器の学習法として用いられている．この手法は，特徴量（弱識別関数）を各集合群やクラスで共有することにより統合的な学習を行うため，効率良く多クラス識別器を構築する．

本論文では，この Joint Boosting をベースとして弱識別器共有をしながら分割統治を行う Divide-and-Conquer Boosting (DaCBoost) を提案する．Joint Boosting は多クラスに対する学習であるためサブカテゴリーをもつ 2 クラス分類には不要な境界を学習してしまう．そこで，DaCBoost では多クラス問題から分割統治アプローチへの適応を行う．図 1 を例にとれば，通常，全体を一括で学習することで得られる「人カテゴリー全体 (a)～(h) に共通する特徴」を共有しながら，「向き，姿勢等サブカテゴリー特有の特徴」を (a, b) や (c, e) 等サブカテゴリーごとに固有学習する．更に，分割統治アプローチではどのように学習サンプルを分割するかが重要である．そこで，Boost 学習で得られる弱識別器に基づく教師なしのクラスタリング手法を提案する．このようなサンプル分割により提案手法は，全てのサンプルに有効な識別器，サブカテゴリー集合に有効な識別器，各サブカテゴリーに特化した識別器の逐次学習を実現する．

2. Joint Boosting

学習アルゴリズムの基礎として用いる Joint Boosting について説明する．一般的な Boosting 手法として AdaBoost [1]，多クラス Boosting として AdaBoost.MH [8] が存在する．AdaBoost をもとに識別性能を向上した学習法として Gentle Boost [6] があり，Joint Boosting はその Gentle Boost を多クラス問題

へ拡張する際，AdaBoost.MH の多クラス学習に特徴の共有という概念を持ち込み，弱識別器数の低減を可能とした学習法である．多クラスに対する Boosting の強識別器，弱識別器はそれぞれ以下となる．

$$H(v,c) = \sum_{m=1}^{M} h_m(v,c) \tag{1}$$

$$h_m(v,c) = \begin{cases} a\delta(v^f > \theta) + b & \text{if } c = c^* \\ k & \text{if otherwise} \end{cases} \tag{2}$$

ここで，v は入力ベクトル，M は弱識別器 h_m の総数，c は識別対象のクラスラベルであり，c^* は弱識別器がどのクラスラベルに対応する弱識別器かを示す添字である．また，v^f はベクトル v の f 番目の要素値を表し，δ 関数は () 内の条件が真である場合 1，偽である場合 0 を応答する．a，b，k の値の具体的な設定はそれぞれ後述する．Joint Boosting は通常の 2 クラス，多クラス Boosting と異なり，クラスの共集合をノード $S(n)$ で表現する．

例えば 3 クラス（ラベル：1，2，3）の場合，ノード $S(n)$ は $S(1):(1)$，$S(2):(2)$，$S(3):(3)$，$S(4):(1,2)$，$S(5):(1,3)$，$S(6):(2,3)$，$S(7):(1,2,3)$ をそれぞれ表現する．これらノードを並列に置き，各ノードについて誤差最小の弱識別器 $h_m(v,S(n))$ を探索する．このように各クラスだけでなく，クラスの共集合もクラスとみなして探索することで，通常の多クラス Boosting ではクラス 1，2 に共に有効な任意の弱識別器 $h(v)$ が $h(v,1)$，$h(v,2)$ として 2 個の弱識別器になることを防ぎ，$h(v,S(4))$ という 1 個の弱識別器で学習される．

ここで，各ノードごとに探索で得られた弱識別器を格納する集合 $G^{S(n)}$ を図 2 のように定義する．弱識別器集合 $G^{S(n)}$ はそれぞれ $S(n)$ の表現するクラス集合に対し識別性能をもつため，図 2 中の上位であるほど複数のクラスで共有可能な識別器となる．つまり，G^{123} は検出対象 1〜3 全てと背景を，G^{12} は検出対象 1 と 2 全てと背景を分離するのに有効な弱識別器群である．この G^{123} と G^{12}（図 2 青領域）は検出対象 1 に対する強識別器 $H(v,1)$（図 2 緑領域），検出対象 2 に対する強識別器 $H(v,2)$（図 2 赤領域）で共有される．この共有により，同様の弱識別器が各クラスごとに学習される無駄を防ぐことができる．各クラスに対する最終的な強識別器はそれぞれ以下となる．

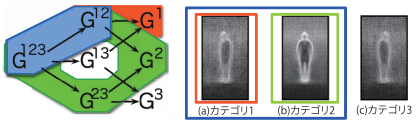

図 2　弱識別器の共有
Fig. 2　Sharing weak-classifier: category (a) and (b) are sharing G^{123}, G^{12}.

$$\begin{aligned} H(v,1) &= G^{123}(v) + G^{12}(v) + G^{13}(v) + G^{1}(v) \\ H(v,2) &= G^{123}(v) + G^{12}(v) + G^{23}(v) + G^{2}(v) \\ H(v,3) &= G^{123}(v) + G^{13}(v) + G^{23}(v) + G^{3}(v) \end{aligned} \tag{3}$$

例えば，c1，c2 共に有効な特徴があればそれは $h_m(v,S(4))$ として発見され，$G^{S(4)}$ に格納される．そして識別時には $H(v,c1)$，$H(v,c2)$ で共に用いられることとなる．

学習は，あるノード $S(n)$ に対する弱識別器 $h(v,S(n))$ を毎ラウンド選択することで進む．その際，最良の弱識別器を選択するために以下の評価関数 J_{wse} という重み付き二乗誤差を用いる．多クラス識別器に対する弱学習の重み付き二乗誤差評価は次のようになる．

$$J_{wse} = \sum_{c=1}^{C}\sum_{i=1}^{N} w_i^c (z_i^c - h_m(v_i,c))^2 \tag{4}$$

C はクラス数，c は各クラスラベルを表し，集合に対する弱識別器 $h_m(v_i,S(n))$ を再度 $h_m(v_i,c)$ として c ごとに表現している．また，z^c は v_i がクラス c に含まれるかどうかを示すラベルであり，$\{1,-1\}$ をとる．各サンプル i のクラス c における重み w_i^c は任意の初期値から $w_i^c = e^{-z_i^c H(v_i,c)}$ として弱識別器が選択されるたびに更新される．式 (4) は識別境界から各サンプルへの二乗距離を指数評価し，その平均値を算出するものである．この評価値を最小にするよう弱識別器を選択することで，汎化性能の高い識別器の構築が期待できる．更に，以下の拡張により弱識別器をノードによる多クラス識別に適応させている．

$$h_m(v,c) = \begin{cases} a\delta(v^f > \theta) + b & \text{if } c \in S(n) \\ k^c & \text{otherwise} \end{cases} \tag{5}$$

これにより各ノードを対象とした識別器を評価できる．ここで，識別器と同様に弱識別器 $h_m(v_i,S(n))$ について評価する際，例えば $h_m(v_i,S(4))$ であれば $h_m(v_i,1)$，$h_m(v_i,2)$ として同時に扱われる点に注意

が必要である．ここで，a，b，k^c のパラメータはそれぞれ以下となる．

$$b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(v_i^f \leq \theta)}{\sum_{c \in S(n)} \sum_i w_i^c \delta(v_i^f \leq \theta)} \tag{6}$$

$$a + b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(v_i^f > \theta)}{\sum_{c \in S(n)} \sum_i w_i^c \delta(v_i^f > \theta)} \tag{7}$$

$$k^c = \sum_{c \notin S(n)} \frac{\sum_i w_i^c z_i}{\sum_i w_i^c} \tag{8}$$

b はしきい値以下の区間における確率密度を，$a+b$ はしきい値以上の区間における確率密度をそれぞれ反映し，負例が多い区間では負方向，正例が多い区間では正方向の強度を示す．k^c は弱識別器 h_m が対象としないクラスについて識別器は k^c を応答として返し，識別関数は介在しない．

3. Boosting に基づく分割統治的戦略による高精度な識別器

提案手法の識別器構築手順は基本的に以下の 2 ステップに分かれる．

- サンプル分割
- サブカテゴリー群に対する分割統治的学習

まず最初に全ての学習データから Gentle Boost による 2 クラス識別器の学習を行う．次に，Gentle Boost で選ばれた弱識別器の分離性能が低下し，これ以上の性能向上の見込みが低くなると，それまで得られた弱識別器の応答を用いてサンプルを分割する（図 3 (a)）．図中の各点群は，図 1 の人画像のような様々な偏りをもつポジティブサンプルと背景画像のようなネガティブサンプルである．この際の尺度には，Gentle Boost 学習時における重み付き二乗誤差の変化量を用い，適当なしきい値 ϵ を設ける．変化量がしきい値を下回ればそれは学習による改善効果が微小であるといえる．本論文では経験的に $\epsilon = 0.01$ と設定した．その後，分割されたサンプルとそれらの集合をノードとして分割統治的な学習を行う（図 3 (b)）．この時点で得られている 2 クラス識別器は全体の識別に対して有効性をもつため，全サンプルを対象とするノード（例．図 2 中の G^{123}）として全サブカテゴリーに利用し，学習を継続する．このような過程で，全体を大まかに識別する識別器，詳細な集合に対して識別を行う識別器をそれぞれ構築していく．こうして得られた各サブカテゴリーへの識別器を統合することで，最終的な識別器が構築される．

3.1 弱識別器応答値を用いた Spectral Clustering による学習サンプルのクラスタリング

データのクラスタリングには通常，特徴量による k-means クラスタリングなどが用いられている．しかし，特徴空間における分布が複雑である場合その特徴空間におけるデータの遠近は識別における有効性と直接的に相関しない．例えば，微小な距離差であっても識別においては非常に有効であることが考えられる．村井らは，Boosting を用いた人領域のセグメンテーションにおいて，類似度の評価に Real AdaBoost による弱識別器の応答値を用いることで，識別器にマッ

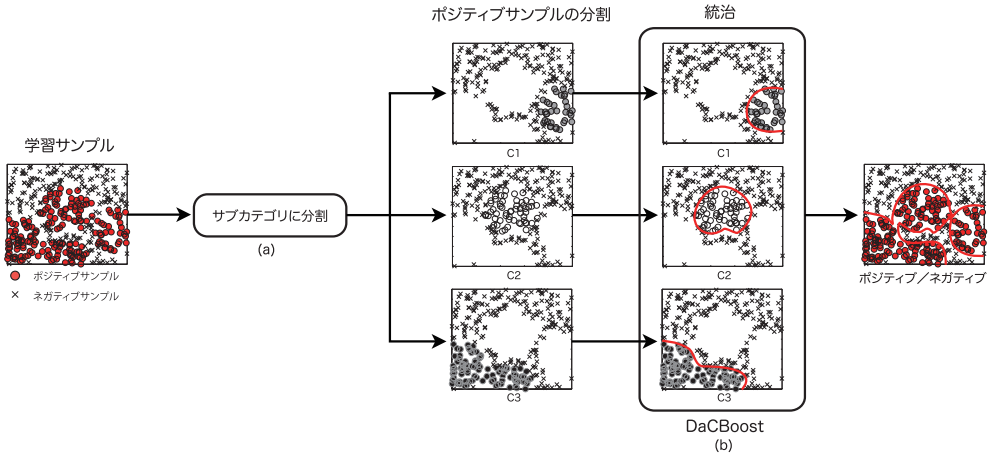

図 3　分割統治的戦略による高精度な識別器
Fig. 3　Constructing a high-accuracy classifier using Divide-and-Conquer Strategy.

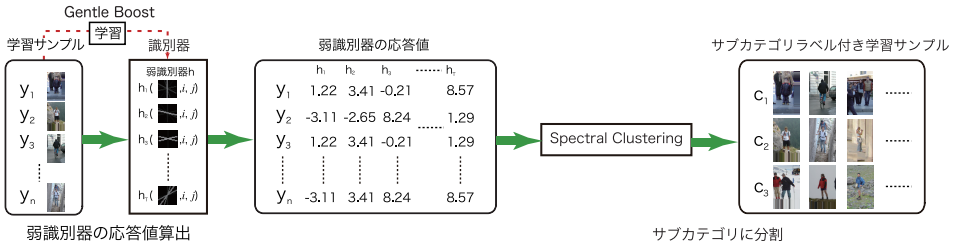

図 4　弱識別器応答値を用いた Spectral Clustering による学習サンプルの分類
Fig. 4　Sample-categorization based on Spectral Clustering using Divide and Conquer Strategy.

チした類似度を算出している [9]．弱識別器を単体で見ると，識別に有効な一次元の空間への写像として考えることができる．この写像される空間は，特徴空間とは異なり識別に有効でない距離は無視されるため，この空間においてクラスタリングすることで識別に必要な分割を得ることができる．そこで，提案手法では図 4 に示すようにポジティブサンプルをクラスタリングする．まず学習サンプルに対して Gentle Boost で学習を行い，弱識別器群を得る．次に，弱識別器の応答値を属性としてサンプルのクラスタリングを行う．これにより，特徴空間での関係性だけでなく，学習を経た識別器の分類情報を用いることができ，より識別に適したクラスタリングとなることが期待できる．

3.1.1　Spectral Clustering

Real AdaBoost や Gentle Boost の弱識別器の応答値は，サンプルが十分でない場合に $\sum_i w_i$ が極小化し，極大化することがあるため，k-means によるクラスタリングではこれらのような応答値を扱う際に危険である．ほかにも，k-means に基づくクラスタリングは，以下の二つの前提に基づいている．一つは，各クラスタのデータ分布は超球形であること，二つ目は，各クラスタの半径は同じ長さであるということである．これらの性質に実データの分布が従うことはまれであり，問題となり得る．そこで，各サンプルの相関行列から次元圧縮を行い，圧縮された空間上でクラスタリングを行う Spectral Clustering [10] を用いる．Spectral Clustering とはサンプル間の相関行列（類似度行列）を作成し，相関行列を固有値問題に帰結し解くことで次元圧縮効果を得る．まず，クラスタリング対象となるサンプルから相関行列 A を求める．相関行列 A は式 (14) で表される

$$A_{ij} = distance(i, j) \tag{9}$$

i と j はクラスタリング対象となるサンプルである．距離関数 distance としては L2 ノルムや内積空間への写像などが用いられる．なお，本論文では L2 ノルムを用いた．この相関行列を，固有値問題に帰結して解く．まず，求められた A を用いて，距離行列 $L = D^{-1/2}AD^{-1/2}$ を求める．ここで，D は $D_{ii} = \sum_{j=1}^{N} A_{ij}$ で求められる対角行列である．

次に，$E = [e_1, \cdots, e_K]$ を計算する．e_k は L の最小固有ベクトルから k 番目の固有ベクトルを抽出したものである．そして，k-means クラスタリングで E をもとにクラスタリングを行う．提案手法では，ここで k-means クラスタリングに代えノンパラメトリックなクラスタリング手法である medoidshifts [11] を用いる．同じくノンパラメトリックなクラスタリング手法である mean-shift [12] において困難である，ファン状のような複雑な分布に対して正しくクラスタリングできる手法である．

3.1.2　提案手法で生成されるサブカテゴリー

人手による 3 クラス（正面/背面，左，右）へのラベリングと提案手法でのラベリングを比較する．本実験では，INRIA Person Dataset [13] を用いる．

(a)　人手によるラベリング

サンプルを分割する際の比較手法として，人手によるラベリングを行う．人手によるラベリングは，人の知識により分類されたサブカテゴリーが得られるため，有効なサブカテゴリー分類法である．しかし，特に多量のサンプルに対しては時間やコストの面に問題がある．

本実験では，INRIA Person Dataset 内の人画像に対して（正面/背面，左，右）へのラベリングを行った．その平均こう配画像を図 5 に示す．図 5 より，それぞれ方向別に分類されていることが分かる．

(b)　弱識別器の応答値によるクラスタリング（提案手法）

提案手法によるクラスタリング結果を図 6 に示す．

図 5　人手によるサンプル分割
Fig. 5　Sample-categorization with user.

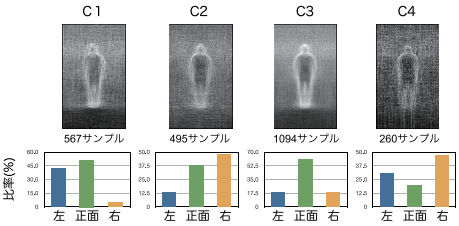

図 6　提案手法によるサンプル分割
Fig. 6　Sample-categorization with propose method.

提案手法では，弱識別器の応答値を入力ベクトルとして Spectral Clustering を行う際，分割するカテゴリー数を自動決定する medoidshifts を組み合わせている．自動分割した結果 4 カテゴリーとなり，前述の人画像の手動分類時とは異なる 4 カテゴリー (C1〜C4) を対象とする．図 6 の平均こう配画像群に着目すると，カテゴリー C1 は頭部形状等にやや左に傾いた傾向が，C2 はやや右に傾いた傾向がある．また，C3 は最も多くのサンプルを含むため，偏りのない滑らかな平均画像が得られている．更に C4 は最もサンプル数が少なく，粗い平均画像である．C4 の特徴として，特に足下の部分にエッジが大きく出ている点と，両足がやや開いていることが挙げられる．

次に，図 6 下段に各カテゴリーに含まれる画像の，方向別の比率を示す．C1，C2，の比率はそれぞれ正面サンプルと左，右方向どちらかを中心としており，人の傾きを中心とした分類を行っていることが分かる．また，C3 は明確に正面を捉えるカテゴリーであることが分かる．C4 はサンプル数自体が非常に少なく，方向としても事前の知見である人の方向とは合致しない傾向をもつ．そこでカテゴリー内の画像を調査したところ，自転車に乗った人などの，足下に複雑なエッジ構造をもつ画像が多く含まれていた（図 7）．これは，本実験における事前の知見にはなかった，「自転車と人」を中心としたカテゴリーを自動的に発見したといえる．

3.2　Divide-and-Conquer Boosting

サブカテゴリーに分類されたデータに対し，識別器を構築する．手法として，Joint Boosting をベースとして 2 クラス問題の分割統治に向け変更を加える．

図 7　c4 カテゴリーのサンプル例
Fig. 7　Example of category c4: Discovered category (c4) means “Human and bicycle”.

3.2.1　Joint Boosting アルゴリズムの分割統治的戦略への適用

Joint Boosting は各学習ラウンドにおいて全ての対象クラスの組合せを探索し，最良の識別器を取得することにより，全クラスで合計した特徴数（弱識別器数）を削減することが可能である．しかし，提案する分割統治戦略を採用する場合，多クラス識別で得られるサブカテゴリー間の弱識別器には多くの無駄が存在する．多クラス識別のための Joint Boosting を，2 クラス識別の分割統治へ適用させるため，各サブカテゴリーにおける弱識別器評価を考える．Joint Boosting では，他のサブカテゴリーを自己と捉えてもエラーと判定するため，2 クラス識別にはあまり寄与しない識別関数を学習してしまう．そこで，各サブカテゴリーごとに存在するサブカテゴリーラベル $z = \{1, -1\}$ を，サブカテゴリーに属するポジティブサンプル，属さないポジティブサンプル，ネガティブサンプルとして $z = \{1, 0, -1\}$ に拡張する．次に，エラー評価を各ノードについて独立に求める．ここでは，他のノードに属するデータを自己と捉えてもエラーと判定せず，同様にそれらを他クラスと捉えても正解と判定しない．

次に，各ノードにおいて下記の評価関数を用いて弱識別器評価を行う．

$$J_{wse} = \sum_{c=1}^{C} \sum_{i=1}^{N} z_i^2 w_i^c (z_i^c - h_m(v_i, c))^2 \qquad (10)$$

上式により，対象ノード内に属さないデータは計算に関与しない．ここで，a，b，k^c のパラメータはそれぞれ Joint Boosting と同様の形で以下の式で計算される．

$$b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(v_i^f \leq \theta)}{\sum_{c \in S(n)} \sum_i z_i^2 w_i^c \delta(v_i^f \leq \theta)} \tag{11}$$

$$a + b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(v_i^f > \theta)}{\sum_{c \in S(n)} \sum_i z_i^2 w_i^c \delta(v_i^f > \theta)} \tag{12}$$

$$k^c = \frac{\sum_i w_i^c z_i}{\sum_i z_i^2 w_i^c} \quad c \notin S(n) \tag{13}$$

これらの操作は，ポジティブサンプルを各サブカテゴリーで独立して保持することと同義である．これにより，各サブカテゴリー間の分離に着目した学習が行われなくなる．また，弱識別器として Gentle Boost と同様の弱識別器を用いているが，本フレームワークは広範に Boosting をベースとした手法に適用できる．

3.2.2 通常の Joint Boosting との相違点

通常の Joint Boosting では，多クラス問題であるため各カテゴリー間を分離するよう重点的に学習される．しかし，本論文で提案するようなサブカテゴリーへの分割統治を用いた 2 クラス問題において，各サブカテゴリー間を分離する境界を学習するのは重要ではない．そこで，提案する DaCBoost では，あるノードについて学習する際そのノードに含まれないサブカテゴリーについて考慮しないことで無駄な境界を学習せず，効率良くサブカテゴリー群とネガティブサンプル間の境界を学習可能である．各サブカテゴリーにおける境界と，統合後の識別境界の例を図 8 に示す．多クラス分類である Joint Boosting は各カテゴリー間の分離を重視しそれぞれの境界を設定する．そのため，c2–c3–ネガティブ間の空間では境界がポジティブサンプルとネガティブサンプルの中央ではなく，ネガティブ判定領域（濃灰色）が広くなり，ポジティブ–ネガティブの 2 クラス識別に適さない．一方，分割統治のためにサブカテゴリー間での識別を考慮しないことで DaCBoost はポジティブサンプルとネガティブサンプルの中央付近に境界が引かれ，2 クラス分類により適した形であることが分かる．

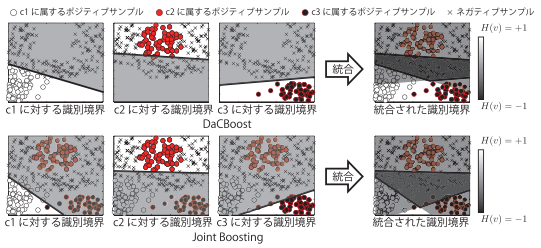

図 8 Joint Boosting と DaCBoost による境界の差
Fig. 8 Difference between DaCBoost and Joint Boosting.

4. 実　　験

提案手法の有効性を示すための実験として，以下の 2 種類を行う．一つは，2 クラス問題に対する識別性能の比較である．ここでは，単体の 2 クラス識別器と，Viola ら [4] のように識別器を複数用いた際の識別性能と比較を行う．この実験を通じて，提案手法による分割統治的アプローチが問題の識別に有益であることを示す．二つ目は，サンプル分割を行う際の分割手法についての比較である．本実験を通じて，弱識別器の応答値を用いてクラスタリングを行うことが提案する分割統治的識別のために有効であることを示す．

4.1 実験概要

本実験では人，車輌それぞれ 1 種のデータセットを用いて識別精度を評価する．各データセット，特徴量，精度の評価基準について述べる．

4.1.1 データセット

実験に使用する各データセットについて説明する．

(a) INRIA Person Dataset [13]

人検出問題として，INRIA から提供されているデータベースを用いる．学習用人画像 2,416 枚と学習用背景画像 1,218 枚，評価用人画像 1,135 枚と評価用背景画像 453 枚が用意されている．

(b) UIUC Image Database for Car Detection [14]

車両検出問題として，UIUC の車両データベースを用いる．道路画像より切り出された車両画像 550 枚と同じく切り出された背景画像 550 枚を学習用に，評価用には道路画像 170 枚をもつデータセットである．

4.1.2 特　徴　量

本論文では，特徴量として Histograms of Oriented Gradients (HOG) [15] を用いる．HOG 特徴量は，局所領域における輝度のこう配方向をヒストグラム化した特徴量であり，ある一定領域に対する特徴量の記述を行う．そのため，大まかな物体形状を表現することが可能であり，人検出や車検出等の物体検出に有効であることが報告されている [2], [3]．

4.1.3 評 価 基 準

横軸に誤検出率，縦軸に検出率を表したグラフによって評価を行う．識別器のしきい値を変化させることによって，汎化性能を考慮した誤識別率に対する識別率の実用的な比較を行うことが可能である．検出率

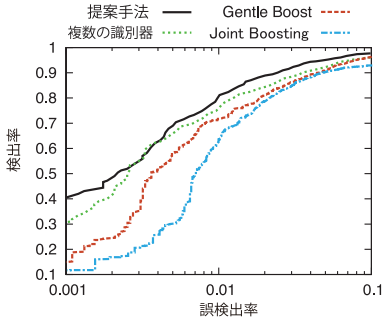

図 9　INRIA Person Dataset での性能比較
Fig. 9　Detection performance on INRIA Person Dataset.

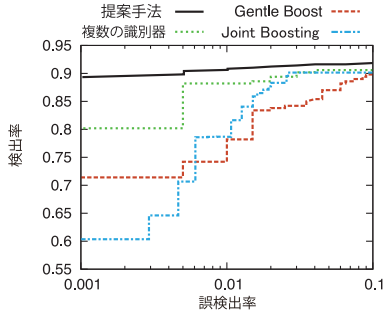

図 10　UIUC Image Database for Car Detection での性能比較
Fig. 10　Detection performance on UIUC Image Database for Car Detection.

と誤検出率は以下のように求める．

$$\text{検出率} = \frac{\text{検出した正解数}}{\text{検出対象数}} \tag{14}$$

$$\text{誤検出率} = \frac{\text{検出した不正解数}}{\text{非検出対象数}} \tag{15}$$

4.2　2 クラス識別器，従来の分割統治との比較実験

提案手法の有効性を示すため，2 クラス分類器である Gentle Boost，それをカテゴリーごとに複数構築した分割統治による識別器，多クラス分類器である Joint Boosting との比較を行う．取り扱う問題は 2 クラス識別としてサブカテゴリー相互の誤りは誤検出としない．

4.2.1　INRIA Person Dataset

図 9 に，INRIA Person Dataset における比較結果を示す．提案手法は先行して学習した Gentle Boost の弱識別器 280 個を用いてサブカテゴリー分割を行っている．全体に，提案手法が Gentle Boost を上回っており，その差は最大で 0.1 以上の向上を見せた．提案手法によるサンプルの分割が有効に働いたことが分かる．また，複数の識別器を用いるよりも提案手法は性能が良く，共有によって効率的に学習が行われたといえる．Joint Boosting はマルチクラス学習であるため，提案手法と同様の分割であるが性能は 2 クラス識別器よりも低下した．誤検出 0.01 付近での急落は，他のサブカテゴリーへの誤解を抑えるための境界が作用して，サブカテゴリー間の境界の出力値が小さくなるためと考えられる．

4.2.2　UIUC Image Database for Car Detection

図 10 に，UIUC Image Database for Car Detection での車両検出における比較結果を示す．提案手法は先行して学習した Gentle Boost の弱識別器 190 個を用いてサブカテゴリー分割を行っている．全体に，提案手法が Gentle Boost を上回っており，その差は最大で 0.1 以上の向上を見せた．車両のように方向などによって大きくエッジの構成が異なる対象に対して，提案手法によるサンプルの分割が有効に働いたと考えられる．また，こちらにおいても提案手法の識別性能は各カテゴリーごとに識別器を構築したものに比べて高く，他クラス識別をそのまま適用した Joint Boosting では分割統治，2 クラス分類に及ばない．

4.3　サブカテゴリー生成手法の比較

提案手法は DaCBoost による実験の際，弱識別器の応答を用いることでより識別に有効な分離を実現する．それを確認するため，人手による 3 クラス（正面/背面，左，右）への分類，弱識別器で選択された特徴量による分類，ランダムなラベリングとそれぞれ比較する．本実験では，INRIA Person Dataset を用いる．

4.3.1　サブカテゴリー生成手法

各比較手法の詳細について以下に示す．

(a)　人手による分類

サンプルを分割する際の比較手法として，人手による分類を行う．本実験では，INRIA Person Dataset 内の人画像に対して（正面/背面，左，右）への分類を行った．**3.1.2** で述べたとおり，その平均こう配画像は図 5 のようにそれぞれ方向別に分類されている．

(b)　ランダムラベリング

サンプルに対するラベル付けをランダムで行う．カテゴリー数のみ手動による方向別分類と同じ 3 として指定した．

(c)　特徴量による分類

一般に，画像データの自動分類は画像より得られた特徴量を用いて行う．しかし，特徴量の距離を使った

表 1　サンプル分割手法の比較 (Detection equal-precision-recall rate)
Table 1　Categorize-method comparison (Detection equal-precision-recall rate).

	サンプル分割手法	検出率 (%)
(a)	人手による分類	81.6
(b)	ランダムラベリング	81.1
(c)	特徴量による分類	81.5
(d)	提案手法	85.5

分類は識別において有利に働くか否かは不明である．ここでは HOG 特徴量全ベクトルを入力として，**3.1.1** で示した分類法で 3 カテゴリーを生成した．

(d) 提案手法

弱識別器の応答値による分類であり，**3.1.2** で述べたとおり結果は図 6 である．人画像を（左向き/正面，右向き/正面，正面，自転車等）として 4 カテゴリーを生成した．

4.3.2 比較結果

INRIA Person Dataset を用いて，Equal-Precision-Recall Rate（EPR：再現率と適合率を等しく設定）時の検出率について比較を行った結果を表 1 に示す．表 1 より，提案手法が最も検出性能が高いことが分かる．また，手動で分類したものと特徴量による分類はそれぞれ同程度であった．これは，方向別などの人間の知識による分類や単純な特徴量による分類は必ずしも識別器に対して最適でないことを示している．提案手法は人の向きという事前知識を用いてはいないが，右向き，左向きなどを考慮しつつ更に識別に有効なカテゴリーとして自転車に搭乗した人というカテゴリーを生成した．更に，これまで正面ラベルを振られていたサンプルのうち相当数が，判別の観点からは右向きや左向きに非常に近しいものであったことが推測される．

5. むすび

人や顔，車両などの向き，姿勢により様々な変動を含む物体を検出するため，サンプルデータ群の偏りをサブカテゴリーとみなして分割し，それらに対して個別に問題を解き，その結果を統合する分割統治による識別器の構築法を提案した．本提案は他の Boosting 手法に容易に展開可能であり，各手法において弱識別器数の低減，識別性能の向上を期待できる．

本論文の貢献は，偏りをもつサンプルを分割する教師なしの手法として弱識別器の応答によるクラスタリング，DaCBoost により 2 クラス識別問題に対する分割統治の実現の 2 点である．人，車両検出問題における評価実験の結果，学習サンプルを分割しない従来の Gentle Boost に比べ I 人検出実験では約 8%（誤識別率 1%時），車輌検出実験において 9.5%（誤識別率 0.1%時）性能を向上させることができ，提案手法の有効性を確認した．また，サンプル分割手法についても特徴量や手動によるラベリングに比べ，弱識別器の応答を用いる提案手法は 3%以上高精度であった．

今後は，より効率的な学習への展開と，ブートストラップなどを組み合わせた総合的な識別への展開についての検討を行う予定である．

文　献

[1] R.E. Schapire and Y. Singer, “Improved boosting algorithms using confidence-rated predictions,” Mach. Learn., no.37, pp.297–336, 1999.

[2] Y. Yamauchi, H. Fujiyoshi, Y. Iwahori, and T. Kanade, “People detection based on co-occurrence of appearance and spatio-temporal features,” National Institute of Informatics Transactions on Progress in Informatics, no.7, pp.33–42, 2010.

[3] 高木雅成，山内悠嗣，藤吉弘亘，“Geometric Context を用いた特徴量間の共起による物体検出の高精度化，” 信学論（D），vol.J93-D, no.8, pp.1429–1438, Aug. 2010.

[4] P.A. Viola and M.J. Jones, “Robust real-time face detection,” Int. J. Comput. Vis., vol.57, no.2, pp.137–154, 2004.

[5] C. Huang, H. Ai, Y. Li, and S. Lao, “Vector boosting for rotation invariant multi-view face detection,” ICCV, pp.446–453, 2005.

[6] J. Friedman, T. Hastie, and R. Tibshirani, “Additive logistic regression: A statistical view of boosting,” Annals of statistics, vol.28, no.2, pp.337–374, 2000.

[7] A. Torralba, K.P. Murphy, and W.T. Freeman, “Sharing features: Efficient boosting procedures for multiclass object detection,” CVPR, pp.762–769, 2004.

[8] Y. Freund and R.E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” J. Comput. Syst. Sci., vol.55, no.1, pp.119–139, 1997.

[9] 村井陽介，藤吉弘亘，伊藤誠也，数井誠人，“弱識別器の応答に基づく類似シルエット画像選択による Chamfer Matching を用いた人領域のセグメンテーション，” 第 12 回画像の認識・理解シンポジウム (MIRU2009)，pp.1468–1475, 2009.

[10] A.Y. Ng, M.I. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” NIPS, 2004.

[11] Y. Sheikh, E. Khan, and T. Kanade, “Mode-seeking by medoidshifts,” ICCV, 2007.

[12] K. Fukunaga and L. Hostetler, “The estimation of the gradient of a density function with application in pattern recognition,” IEEE Trans. Inf. Theory, vol.21,

Issue 1, pp.32–40, 1975.
[13] INRIA Person Dataset,
http://ralyx.inria.fr/2006/index.html
[14] UIUC Image Database for Car Detection,
http://l2r.cs.uiuc.edu/cogcomp/Data/Car/
[15] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” CVPR, vol.1, pp.886–893, 2005.

（平成 23 年 6 月 6 日受付，10 月 3 日再受付）

土屋　成光　（正員）

2008 中部大学大学院博士前期課程了．2011 同博士課程満期退学．2011 中部大学研究員．パターン認識・理解，機械学習の研究に従事．

藤吉　弘亘　（正員）

1997 中部大学大学院博士後期課程了．1997〜2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow．2000 中部大学講師，2004 同大准教授を経て 2010 より同大教授．2005〜2006 米カーネギーメロン大学ロボット工学研究所客員研究員．計算機視覚，動画像処理，パターン認識・理解の研究に従事．2005 年度ロボカップ研究賞．2009 年度情報処理学会論文誌コンピュータビジョンとイメージメディア優秀論文賞，2009 年度山下記念研究賞．情報処理学会，電気学会，IEEE 各会員．