

第3章

分割統治戦略による識別器の高精度化

本章では、従来の物体検出に用いられる統計的学習法に分割統治戦略を導入することで高精度化するアプローチについて提案する。まず、3.1.1にて複雑な変化を持つ物体を分類する従来法について述べた後、3.1.2にて提案手法である分割統治戦略の導入について述べる。そして、3.2では分割過程における具体的なサンプル分割法として弱識別器の応答値に基づくサンプルの分割について述べた後、3.3で分割した学習サンプルに対する効率的な学習法として Divide-and-Conquer Boosting について述べる。

3.1 複雑な物体分類の問題と分割統治戦略の導入

3.1.1 複雑な変化を持つ物体の分類

2.4.1 では、物体の複雑さにより識別性能が低下する問題を、図 2.7 を例として述べた。図 2.7 は人検出用データベースとして公開されている INRIA person dataset[3] に含まれている画像である。これらの画像は人が中心となるように配置されているが、体幹の傾き、顕著な年齢差、複雑な姿勢等変化が含まれている。これらの変化を吸収するように識別することは非常に困難な問題である。

従来、このような変化を伴う検出対象に対する解決策としては Viola らの顔検出 [15] に習うことが多い。この手法は Haar-like 特徴量と Boosting を用いて、顔の向きの変化に対する頑健性を得るために、学習画像を顔の傾きや方向毎に人手で分割して、サブカテゴリ毎に個別の識別器を構築するアプローチである。方向、傾き毎に構築した識別器のどれか一つでも反応すれば顔と判定する。また、Huang ら [20] は Viola らと同様に顔画像を右、正面、左の 3 クラスに手で分類し、各向きを Vector Boosting というマルチ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。しかし、これらの手法では、人手によりサブカテゴリに分類してラベル付けするためのコストが非常に大きい。さらに、構築される識別器の性能は、サブカテゴリ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 2.7(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。このように、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 2.7(f) は正面を向いているため頭、肩等においては図 2.7(a), (b) と同じ特徴を持つといえる。また、同時に、体幹が傾きにより脚部が左下方向に伸びている点で左向きの画像である図 2.7(c) と同じ特徴を持つといえる。識別においても、サブカテゴリ毎に分割して構築された識別器は「サブカテゴリ (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 (顔, 人など) らしさ」について同様の情報をそれぞれ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識別時において計算コストを引き上げる原因となる。

3.1.2 分割統治戦略の導入

3.1.1で述べたように、統計的学習法による2クラス分類では、同一検出対象内に多様性が存在すると問題が複雑化する。そこで本研究では、まず検出対象となる学習サンプルを自動的にサブカテゴリに分割する手法を提案する。次に、サブカテゴリ固有の情報は固有に、カテゴリ全体やカテゴリ群に共通する情報は共有して識別器を構築する手法を提案し、両手法の組み合わせによる分割統治アプローチを提案する。

本研究では、Joint Boosting をベースとして、弱識別器を共有しながら分割統治を行う Divide-and-Conquer Boosting (DaCBoost) を提案する。Joint Boosting はマルチクラスに対する学習であるため、サブカテゴリを持つ2クラス分類には不要な境界を学習することがある。そこで、提案する DaCBoost ではサブカテゴリ同士の分類を重視せず、各サブカテゴリと背景の境界のみを学習し、それらの結果を統合することで高精度な2クラス識別器を実現する。図3.1に提案手法の流れを示す。

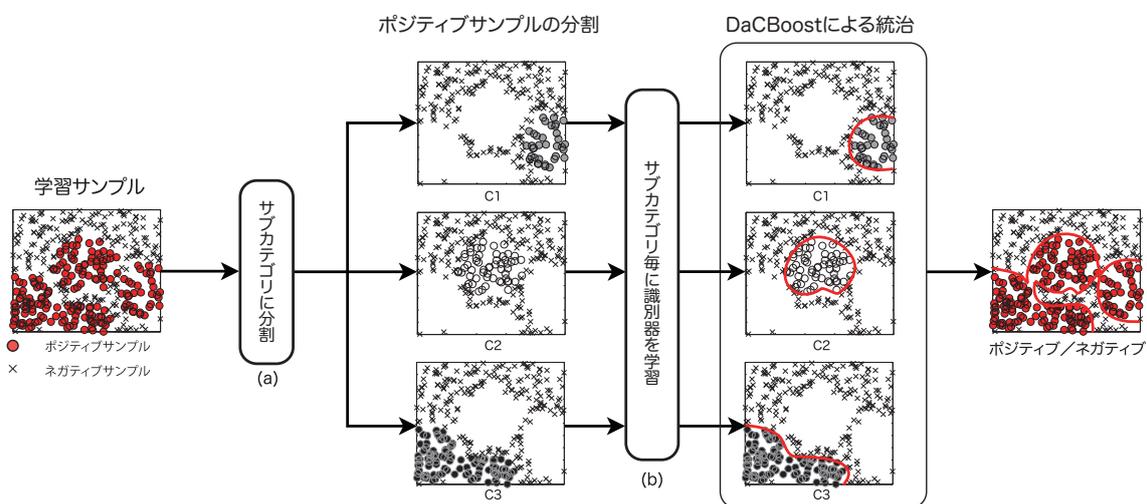


図 3.1: 分割統治的戦略による識別器の学習法 (概念図)。

提案手法の分割統治識別器を構築する手順は以下の2ステップとなる。

≤ サンプル分割

まず最初に、学習サンプルの全てのデータに対して1個の2クラス分類器を学習する。これは、通常の gentleboost による2クラス分類と同一の学習である。次に、gentleboost で得られた弱識別器の応答を用いて、図3.1(a)のようにサンプルを分割する。

≤ 共有学習

分割されたサンプル集合をサブカテゴリとし、図2.6で示した Joint Boosting と同様にサブカテゴリとその組み合わせをノードとして共有学習を行う。その結果、図3.1(b)のように各サブカテゴリに対し識別器が構築される。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

対して有効性を持つため、全サンプルを対象とするノード（例、図 2.6 中の G^{123} ）として全サブカテゴリに利用し、学習を継続する。

以上の 2 つのステップにより、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する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への識別器を統合することで、最終的な強識別器を構築する。

3.2 弱識別器の応答値に基づくサンプルの分割

従来、データを分割する場合には、特徴量空間における分布を利用して k-means などのクラスタリングを行う。しかし、特徴量空間は、識別対象のサンプルとそれ以外の背景サンプルの識別に寄与しない冗長な特徴次元を含んでいる。学習サンプルデータを分割するためには、識別対象のサンプルとそれ以外の背景サンプルを識別するために有効な特徴量のみを用いて分割することが必要である。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、識別器にマッチした類似度を算出している [21]。本手法は、Boosting の各弱識別器を識別に有効な 1 次元の特徴ベクトルと考えることができると述べている。弱識別器の応答値で構成される空間は、特徴空間とは異なり識別に有効でない特徴量は無視されるため、この空間においてクラスタリングすることで識別に有効な分割を得ることができる。そこで、提案手法では図 3.2 に示すように、まず学習サンプルに対して gentleboost 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情報を用いることができるため、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。

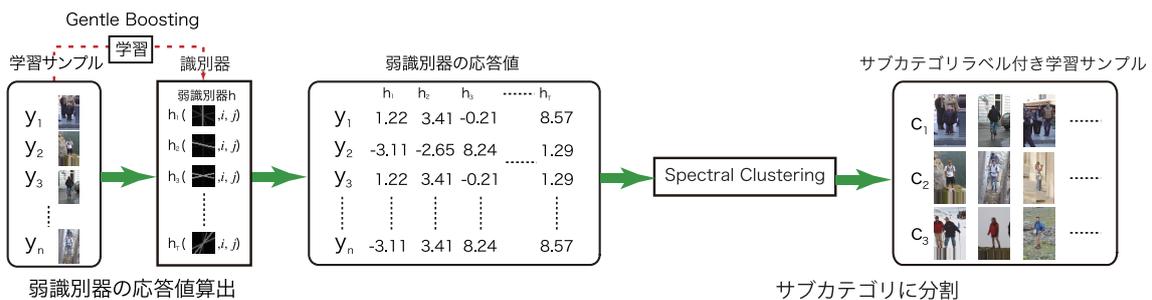


図 3.2: 弱識別器応答値を用いた学習サンプルのカテゴリライズ。

3.2.1 Spectral Clustering

k-means に基づくクラスタリングは、以下の2つの前提に基づいている。1つは、各クラスタのデータ分布は超球形であること、2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稀であり、問題となり得る。そこで、本研究では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[22] を用いる。Spectral Clustering は k-means に比べより多くの計算量を必要とするが、データの連結性などに着目したクラスタリング結果を得ることができる。図 3.3 に Spectral Clustering の流れを示す。

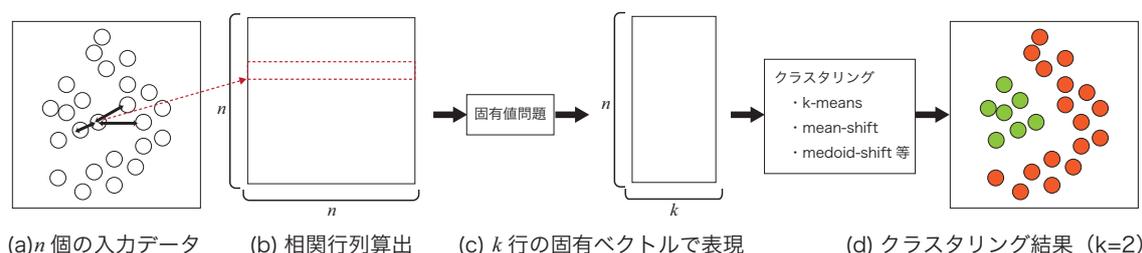


図 3.3: Spectral Clustering の流れ.

Spectral Clustering 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ことができる。まず、図 3.3(a) のようなクラスタリング対象となるサンプルから、図 3.3(b) のように相関行列 A を求める。相関行列 A は式 (3.1) で表される。

$$A_{ij} = distance(i, j) \quad (3.1)$$

ここで、 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、本研究では 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、図 3.3(c) のように固有値 $E = [e_1, \dots, e_k]$ 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、 E を基に k-means クラスタリングすることで、図 3.3(d) のように任意のクラスタ数に分割することが出来る。

3.2.2 ノンパラメトリックなクラスタリングへの対応

提案手法では、クラスタ数をパラメータとして指定しないノンパラメトリックなクラスタリングを行う必要がある。そこで、Spectral Clustering の最後の処理に medoid-shift[23] を用いる。medoid-shift

は、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[24] と類似しているが、図 3.4 に示すように、mean-shift がデータのセントロイド（平均値）を仮定して摂動し、クラスタリングするのに対し、入力されたデータの中央値を各クラスの代表として摂動し、各データ間の相互距離関係によりクラスタリングを行う点が異なる。medoid-shift は、データが存在しない特徴空間の点をクラスを中心とすることはしない。そのため、mean-shift において困難となるファン状のような複雑なデータ分布に対して、比較的正しくクラスタリングできる手法である。

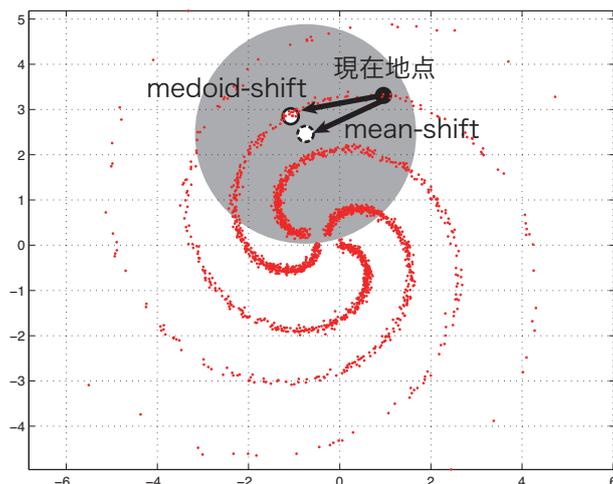


図 3.4: medoid-shift と mean-shift の違い。

3.2.3 サンプル分割結果

ここでは、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。データベースには、INRIA Person Dataset[3] を用いる。以下に人手によるサンプル分割と提案手法におけるサンプル分割について述べる。

人手によるラベリング

人手によるラベリングは、人の知識によりサンプルを分類するサブカテゴリ分類法である。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。また、サンプルにどのラベルを付与するかは作業者の主観に基づくため、作業者によりばらつきが出るという問題もある。

本実験では、INRIA Person Dataset の人画像 2,416 枚に対して、3つのサブカテゴリ（正面／背面，左，右）にラベリングを行った。各カテゴリの枚数は、正面／背面 1,308 枚，左向き 554 枚，右向き 554 枚である。各カテゴリに分類された画像から求めた平均勾配画像と各カテゴリのサンプル例を図 3.5 に示す。図 3.5 より、それぞれ方向別に分類されていることがわかる。

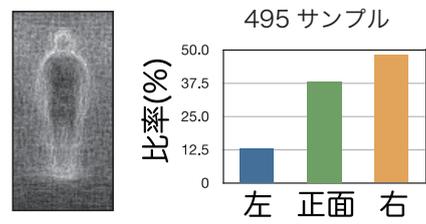
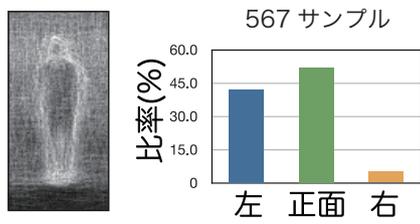


図 3.5: 人手によるサンプル分割.

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の平均匀配画像とサンプル例を図 3.6 に示す. 弱識別器の応答値をクラスタリングした結果, 人画像は 4 つのカテゴリ $C1 \in C4$ に自動的に分割された. 図 3.6 のグラフはそれぞれ $C1 \in C4$ 各カテゴリに対して, 方向別に人手でラベル付けしたサンプルがどの程度含まれていたかを比率で示している. 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, 分割するカテゴリ数を自動的に決定する Medoide-shift を組み合わせている. 図 3.6 の平均匀配画像群に着目すると, カテゴリ $C1$ 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, 偏りの無い滑らかな平均画像が得られている. 一方, $C4$ は最もサンプル数が少ないため, 粗い平均画像である. $C4$ の特徴としては, 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傾向がある.

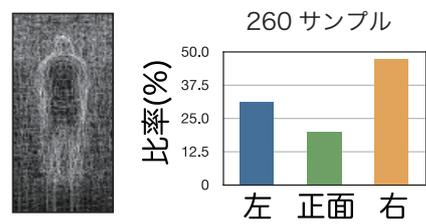
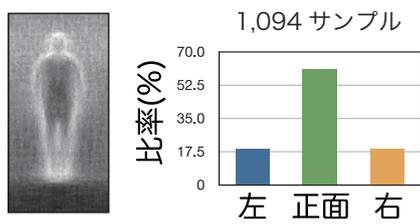
次に, 各カテゴリに含まれるサンプルの方向別の比率と, 各カテゴリのサンプル例に注目する. 図 3.6(a)(b) より, $C1, C2$ の比率はそれぞれ正面/背面サンプルと左, 右方向どちらかを中心としており, 人の傾きを中心としたカテゴリ分けを行っていることがわかる. また, 図 3.6(c) より, $C3$ は正面/背面サンプルを多く含み, 平均匀配画像からも明確に正面/背面を捉えるカテゴリであることがわかる. 図 3.6(d) より $C4$ はサンプル数自体が非常に少なく, 方向としても事前の知見である人の方向とは合致しない傾向を持つ. そこでサンプルを調査したところ, このカテゴリにクラスタリングされたサンプルの中には, 自転車に乗った人が含まれており, 脚部に複雑なエッジ構造を持つことが判明した. これは, 本実験における事前の知見には無かった「自転車に乗っている人」を中心としたカテゴリを自動的に発見したといえる.



(a)C1



(b)C2



(c)C3



(d)C4

図 3.6: 提案手法によるサンプル分割.

3.3 Divide and Conquer Boostingによるサブカテゴリに対する効率的な学習

3.2でサブカテゴリに分割されたデータに対し、各サブカテゴリ毎に識別器を構築する。本手法では、Joint Boostingをベースとしてサブカテゴリを含む2クラス問題向けに適応を行い、各サブカテゴリ間で共通して有効な弱識別器を共有して学習する..

3.3.1 Joint Boosting アルゴリズムの分割統治戦略への適用

Joint Boosting は各学習ラウンドにおいて全ての対象クラスの組み合わせを探索し、最良の識別器を学習することにより、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。本研究では、全サブカテゴリを対象としたものが通常の2クラス問題を解く gentleboost と同様であることに着目し、段階的にサンプルを分割しながら学習を行うことで Joint Boosting と類似した識別器の構造を実現する。

まず、各サブカテゴリにおける弱識別器の評価を考える。Joint Boosting では、カテゴリ間のエラーについても背景カテゴリとのエラー同様に評価するため、サブカテゴリ問題にはあまり寄与しない識別関数についても学習する。そこで、各サブカテゴリ毎に存在するサブカテゴリラベル $z = \{1, 1\}$ を、サブカテゴリに属するポジティブサンプル、属し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の評価を行う。

$$J_{wsc} = \sum_{c=1}^C \sum_{i=1}^N z_i^2 w_i^c (z_i^c - h_m(\mathbf{x}_i, c))^2 \quad (3.2)$$

上式により、対象ノード内に属し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。

$$b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(\mathbf{x}_i^f \sim \theta)}{\sum_{c \in S(n)} \sum_i z_i^2 w_i^c \delta(\mathbf{x}_i^f \sim \theta)} \quad (3.3)$$

$$a + b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(\mathbf{x}_i^f > \theta)}{\sum_{c \in S(n)} \sum_i z_i^2 w_i^c \delta(\mathbf{x}_i^f > \theta)} \quad (3.4)$$

$$k^c = \frac{\sum_i w_i^c z_i}{\sum_i z_i^2 w_i^c} \quad c \notin S(n) \quad (3.5)$$

これらの操作は、ポジティブサンプルを各サブカテゴリで独立して保持することと同義である。これにより、各サブカテゴリ間の分離に着目した学習が行われなくなる。提案する DaCBoost の学習アルゴリズムを **アルゴリズム 3.1** に示す。

アルゴリズム 3.1: DaCBoost の学習アルゴリズム.

1. 入力: I 個の学習サンプル $\{\mathbf{x}_1, z_1 \dots \mathbf{x}_I, z_I$ を用意する
 \mathbf{x} は学習サンプル, $z_i \in \{1, 1, \dots, C\}$ は学習サンプルが属するクラスラベルを表す

2. 初期化: 学習サンプルの重み D を初期化

$$w_1^c(i) = 1/I \quad (3.6)$$

3. 学習:

• C 個のクラスの全ての組み合わせを示す N 個の要素を持つ集合をノード $S(n)$ とする

For $m = 1, \dots, M$ //学習ラウンド

For $n = 1, \dots, N$ //すべてのノード

• 評価値 J_{wse} が最小の弱識別器 $h_m(\mathbf{x}, c)$ を選択

$$h_m(\mathbf{x}, c) = \begin{cases} a\delta(\mathbf{x}^f > \theta) + b & \text{if } c \in S(n) \\ k^c & \text{otherwise} \end{cases} \quad (3.7)$$

• 評価値 J_{wse} の算出

$$J_{wse} = \sum_{c=1}^C \sum_{i=1}^N z_i^2 w_i^c (z_i^c - h_m(\mathbf{x}_i, c))^2 \quad (3.8)$$

• $h_m(\mathbf{x}, c)$ のパラメータ a, b, k^c の算出

$$b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(\mathbf{x}_i^f \sim \theta)}{\sum_{c \in S(n)} \sum_i z_i^2 w_i^c \delta(\mathbf{x}_i^f \sim \theta)} \quad (3.9)$$

$$a + b = \frac{\sum_{c \in S(n)} \sum_i w_i^c z_i^c \delta(\mathbf{x}_i^f > \theta)}{\sum_{c \in S(n)} \sum_i z_i^2 w_i^c \delta(\mathbf{x}_i^f > \theta)} \quad (3.10)$$

$$k^c = \frac{\sum_i w_i^c z_i}{\sum_i z_i^2 w_i^c} \quad c \notin S(n) \quad (3.11)$$

End for

• 学習サンプルの重み w を更新

$$w_{t+1}^c(i) = w_t^c(i) \exp(z_i^c - h_m(\mathbf{x}_i, c))^2 \quad (3.12)$$

End for

4. 出力: 強識別器

$$H(\mathbf{x}, c) = \sum_{m=1}^M h_m(\mathbf{x}, c) \quad (3.13)$$

3.3.2 Joint Boosting との相違点

Joint Boosting では、あるノードについて識別器を評価する際に他のサブカテゴリを負例として用いる。このアルゴリズムにより決定される境界は、各サブカテゴリ間を分離するよう重点的に学習される。しかし、サブカテゴリを含む2クラス問題では、各サブカテゴリは本来同一クラスであるため、これらを分離する境界を学習するのは無駄となる。提案する DaCBoost では、あるノードについて計算するとき、そのノードに含まれないサブカテゴリについては考慮しないため無駄な境界を学習することがない。これにより、効率よくサブカテゴリ群とネガティブサンプル間の境界を学習可能である。図 3.7 に示すように、サブカテゴリ間での識別を考慮しないことで DaCBoost はより多くのポジティブ/ネガティブの分離が可能である。

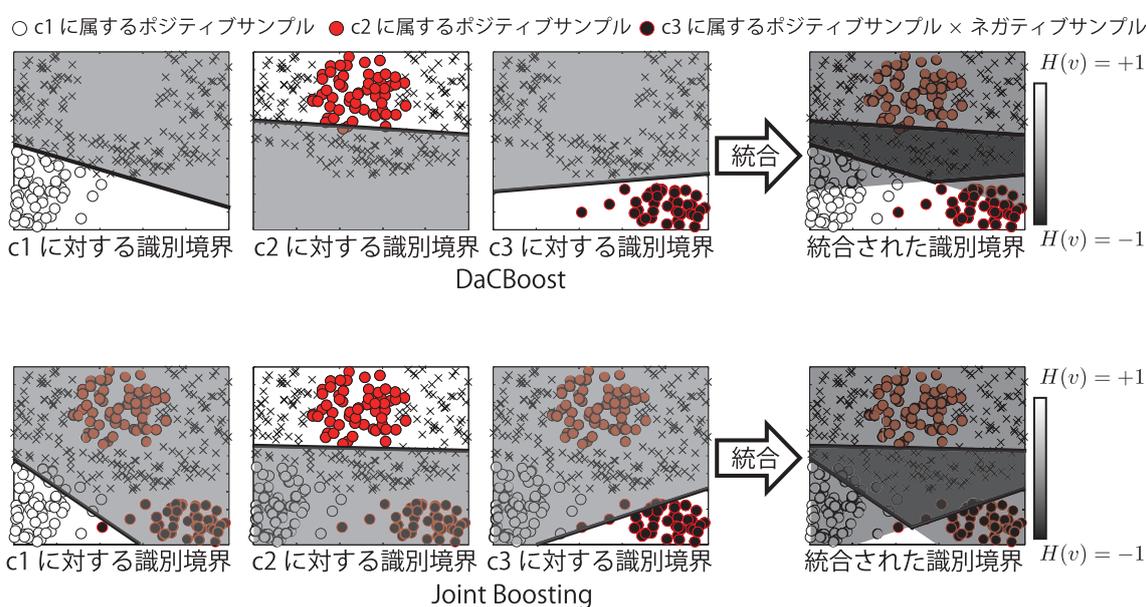


図 3.7: Joint Boosting と DaCBoost による境界の差。

3.4 評価実験

提案手法の有効性を示すために、以下の2種類の実験を行う。1つは、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アプローチによる識別に有効であることを示す。2つ目は、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 らのように識別器を複数用いた際の識別性能との比較を行う。この実験を通じて、提案手法による分割統治アプローチが問題の識別に有益であることを示す。

3.4.1 実験概要

本実験では、人、車両それぞれ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。まず、実験に使用する各データセットについて説明する。

INRIA Person Dataset[3]

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。

UIUC Image Database for Car Detection[25]

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 枚を持つデータセットである。

本実験では、人と車両共に特徴量として Histograms of Oriented Gradients (HOG) [3] を用いる。HOG 特徴量は、局所領域における輝度の勾配方向をヒストグラム化した特徴量として記述する。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 [13][14]。

3.4.2 弱識別器の応答値によるサンプル分割の有効性

提案手法によるサンプル分割は、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。その有効性を確認するために、各手法で分割したサンプルを用いて識別器を構築した際の識別性能を比較する。各比較手法の詳細について以下に示す。

(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。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。3.1.3 で述べた通り、その平均勾配画像は図 3.5 のようにそれぞれ方向別に分類されていることがわかる。

(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。カテゴリ数のみ手動による方向別ラベリングと同じ 3 として指定した。

(c) 特徴量によるクラスタリング

一般に、画像データの自動クラスタリング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カテゴリ分け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用いたクラスタリングでラベリングする。

(d) 提案手法

弱識別器の応答値によるクラスタリングであり、3.1.3 で述べた通り結果は図 3.6 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カテゴリにラベリングを行っている。

■ 実験結果

本実験では、INRIA Person Dataset を用いる。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 3.4.2 に示す。表 3.4.2 より、提案手法である弱識別器の応答値による分割手法が最も検出性能が高いことがわかる。また、手動でラベリングしたものと特徴量によるクラスタリングはそれぞれ同程度である。これは、方向別などの人間の知識によるラベリングや単純な特徴量によるラベリングでは、必ずしも識別器に対して最適では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として自転車に搭乗した人というカテゴリを生成した。更に、人手では正面ラベルを付加されたサンプルの内相当数に、判別の観点からは右向きや左向きに非常に近いものが存在しているといえる。

表 3.1: サンプル分割手法の比較。

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

3.4.3 DaCBoost によるサブカテゴリ学習の有効性

提案手法の有効性を示すため、2クラス分類器である gentleboost、それをカテゴリ毎に複数構築した分割統治による識別器との比較を行う。取り扱う問題は2クラス識別として、サブカテゴリ相互の誤りは誤検出としない。本実験では、各手法の弱識別器の総数を500個と固定とする。各比較手法の詳細について以下に示す。

gentleboost

gentleboost は2クラス識別器であり、サンプル分割を行わない際の識別性能を示す。

複数の識別器

複数の識別器は One vs Other 戦略によりクラス毎に構築した gentleboost 識別器群であり、分割したサンプルに対して弱識別器の共有を行わず学習した際の識別性能を示す。

提案手法 (DaCBoost)

提案手法は DaCBoost により構築した識別器であり、分割したサンプルに対して共通した特徴を共有しながら効率的に学習した際の識別性能を示す。

実験結果の比較には Detection Error Tradeoff(DET) カーブを用いる。DET カーブは横軸に False Positive Per Window(FPPW), 縦軸に miss rate を表わし、左下の原点に近いほど検出性能が高いことを示す。FPPW は、全背景画像に対し人のサイズに合わせたウインドウをラスタスキャンし、人と誤検出したウインドウの割合を表す。miss rate は全ての切り出した人画像に対し検出を行い、未検出であった割合を表す。これらはトレードオフの関係にあり、識別器の出力 $H(\boldsymbol{x})$ を人と判定するしきい値によって上下する。FPPW, miss rate はそれぞれ以下の式で計算される。

$$FPPW = \frac{\text{False positive の数}}{\text{全てのテスト用背景画像}} \quad (3.14)$$

$$\text{miss rate} = \frac{\text{False negative の数}}{\text{全てのテスト用人画像}} \quad (3.15)$$

ここで、False positive は誤って人として検出された背景画像、False negative は人として検出されなかった人画像である。

■ INRIA Person Dataset

図 3.8 に、INRIA Person Dataset における比較結果を示す。全体に、提案手法が gentleboost に比べ低い miss rate を達成しており、FPPW= 0.01 時点で 0.08 向上 (検出率 8% 向上に相当) し、最大で 0.1 以上の性能向上を見せた。提案手法によるサンプルの分割が有効に働いたことがわかる。また、同数の弱識別器でありながら、サブカテゴリ毎に独立して識別器を構築するよりも提案手法は平均して 0.05 miss rate が低く、DaCBoost で特徴を共有しながら、分割されたデータを学習することによって効率的に学習が行われたといえる。

■ UIUC Image Database for Car Detection

図 3.9 に、UIUC Image Database for Car Detection での車両検出における比較結果を示す。全体に、提案手法が gentleboost を miss rate において下回っており、FPPW= 0.001 時点で 0.095 向上 (検出率 9.5% 向上に相当) し、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いえる。また、本データセットにおいても提案手法の識別性能は、各カテゴリ毎に識別器を構築する手法に比べて高い。

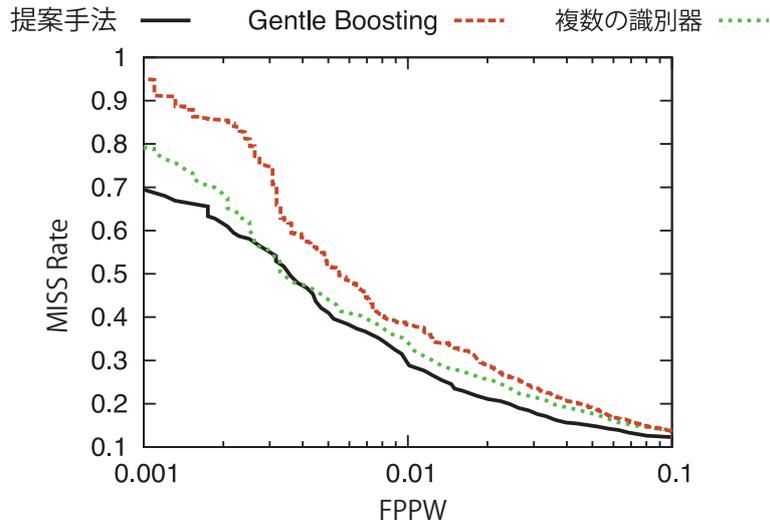


図 3.8: INRIA Person Dataset での性能比較.

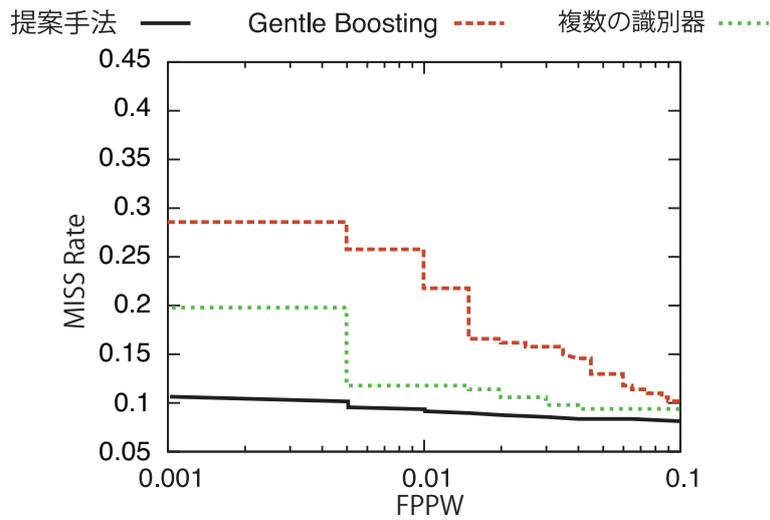


図 3.9: UIUC Image Database for Car Detection での性能比較.

3.5 まとめ

本章では、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と見なして分割し、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アプローチによる識別器の構築法を提案した。

本提案の貢献は、1つ目に偏りを持つサンプルを分割する教師無し的手法として弱識別器の応答

によるクラスタリング, 2つ目は DaCBoost により 2 クラス分類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 gentleboost と比較した結果, INRIA Person Dataset では約 8% (FPPW0.01 時), UIUC Image Database for Car Detection においては約 9.5% (FPPW0.001 時) 性能を向上させることができた。また, サンプル分割手法についても従来の特徴量や手動によるラベリングと比較して, 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。

本提案は 2 クラス問題における Boosting ベースの手法に容易に展開可能であり, 弱識別器数の低減, 識別性能の向上が期待できる。