



相補的な視覚的学習による複数の認識手法の統合

野宮, 浩揮
上原, 邦昭

(Citation)

電子情報通信学会論文誌. D, 情報・システム, J90-D(11):3043-3054

(Issue Date)

2007-11-01

(Resource Type)

journal article

(Version)

Version of Record

(Rights)

Copyright (c) 2007 IEICE

(URL)

<https://hdl.handle.net/20.500.14094/90000829>



相補的な視覚的学習による複数の認識手法の統合

野宮 浩揮^{†a)} 上原 邦昭^{†b)}

Complementary Visual Learning to Integrate Multiple Recognition Approaches

Hiroki NOMIYA^{†a)} and Kuniaki UEHARA^{†b)}

あらまし 本論文では、汎用的で正確な物体認識手法の構築を目的として、機械学習を用いて効果的に画像中の物体を認識するための視覚的学習法を提案する。一般に、実世界に存在する物体は非常に複雑であり、単一の特徴を用いて正しく認識することは困難である。実際、人は様々な特徴を認識に用いている。そこで、様々な特徴に基づく認識手法を効果的に統合するために、相補的学習法を提案する。相補的学習法では、それぞれ異なる特徴を用いた複数の分類器が互いの欠点を補うように学習するため、従来の投票法などの単純な統合手法に比べて、高い認識精度を実現できる。また、相補的学習を用いて統合するための認識手法として、物体の形状と領域特徴に基づく2種類の認識手法を提案し、実世界の物体を用いた認識を行う。更に、多数の認識手法を用いることが可能となるように相補的学習法を拡張し、その有効性を検証する。

キーワード 相補的学習, 機械学習, 形状ベースの認識, 領域ベースの認識, NadaBoost

1. ま え が き

コンピュータビジョンの分野では、物体を正しく認識するために、多くの認識手法が考案されている。しかし、従来のパターン認識の枠組みに基づく認識手法には、正確な認識が可能な対象が限定的であったり、認識課題に応じてモデルやテンプレートなどをあらかじめ定義しておかなければならないといった問題がある。これらの問題を解決し、より汎用的な認識手法を確立するために、視覚的学習と呼ばれる効果的な手法が提案されている[1]~[3]。視覚的学習は、物体の認識に機械学習を利用することにより、人の視覚と同等な物体認識を行うことを目的としている。機械学習の導入により、認識手法の汎用性や認識精度が改善されることが示されているため、近年盛んに研究されている。

しかし、従来の視覚的学習は、二つの主要な問題を抱えている。一つは、その多くが単戦略学習(monos-trategy learning)を行っている点である。画像には、輪郭やテクスチャなど、複数の高レベル特徴が含まれ

ているが、これらの認識手法は、単一の高レベル特徴あるいはそれに属する低レベル特徴(例えば、テクスチャを高レベル特徴とすると、画素の明るさや色相など)しか考慮していない。実世界の物体は非常に複雑であるため、単一の高レベル特徴のみを用いて正確に識別することは困難である。人のもつ優れた視覚も、複数の高レベル特徴を適切に統合することによって成り立っている[4]。つまり、汎用的な認識のためには、多戦略学習(multistrategy learning)を行う必要がある。すなわち、多様な高レベル特徴に基づいて複数の分類器を構成し統合すべきである。多戦略学習法は、アンサンブル学習の観点からみても有効である[5]。

もう一つの問題点は、既存の視覚的学習法の多くが、パラメータ決定などの最適化問題を解くためにのみ機械学習を用いている点である[2],[3]。これに対し、本論文では視覚的学習を人の視覚を表現・記述するための枠組みと考え、人の視覚モデルに基づいた認識を行うことを目的としている。

人の視覚は複数の高レベル特徴を用いていることが既に知られているため、多様な高レベル特徴に基づく認識手法を統合する必要がある。しかし、投票法のような単純な統合法では、大幅な改善は期待できない。そこで、より効果的に多戦略学習を行うために、相補的な視覚的学習法を提案する。これは、複数の分類器

[†] 神戸大学大学院自然科学研究科, 神戸市

Graduate School of Science and Technology, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe-shi, 657-8501 Japan
a) E-mail: nomiya@ai.cs.scitec.kobe-u.ac.jp

b) E-mail: uehara@kobe-u.ac.jp

が協調して学習できる多戦略学習手法である。具体的には、ある分類器が誤分類した事例を、他の分類器で重点的に学習し、分類器同士が互いの欠点を補いながら学習を進めていく。結果的に、各分類器を別々に訓練する単純な統合法よりも効果的な学習が可能となる。

本論文では、学習アルゴリズムの記述を簡潔にするため、まず二つの認識手法を用いる場合の多戦略学習法について述べ、最終的に多数の認識手法を用いるための拡張について述べる。二つの高レベル特徴に基づく物体認識手法は以下の二つからなる。一つは形状に基づいており、輪郭の断片の集合を認識に用いる。輪郭の断片の集合は、視覚心理学における群化の概念[4]に基づいて、意味のある構造にグループ化される手法である。もう一つは領域に基づいており、すべての輪郭片を含む最小の領域内に存在する、画素の階調値の分布を認識に用いる手法である。

2. 関連研究

近年、効率的な物体認識を行うための手法として、様々な視覚的学習法が提案されている。例えば、Maréeらは、決定木のアンサンブルを用いた画像分類手法を提案している[2]。この手法では、画像からランダムに抽出されたサブウィンドウを用いて物体が記述され、多くのサブウィンドウから構成される決定木のアンサンブルを用いて分類が行われる。Tuは、決定木を用いたboostingにより物体を識別する手法を提案している[3]。この手法では、確率的boosting木と呼ばれる、節点が分類器に対応している決定木が用いられている。これらの手法は、モデルなどをあらかじめ定義することなく、多様な認識課題に適用できる。しかし、単戦略学習のため、複雑な物体の認識には限界がある。

これに対し、Nomiyaらは2種類の認識手法を統合する多戦略学習法を提案している[1]。形状ベースと領域ベースの認識手法により生成された分類器を、決定木と判別分析を用いて統合しているが、各分類器は別々に生成され、投票法で単純に統合しているのみで、認識性能が十分改善できているとはいえない。

そのほかの多戦略学習を用いた認識手法として、co-trainingに基づくものがある。co-trainingとは、二つの分類器が互いに相手の学習結果を考慮して、繰り返し学習を行う手法であり、半教師付き学習に用いられる。学習を行う際には、あらかじめ分類された画像が訓練データとして必要になるが、人手で多くの画像を分類するのは困難である。しかし、分類されてい

ない画像は、多量に収集するのが容易なので、半教師付き学習法は物体認識に有効であるといわれている。Levinらは、co-trainingに基づいて多戦略学習を構成し、物体の検出を行う手法を提案している[6]。この手法では、二つの分類器がそれぞれあらかじめ分類されていない訓練事例を分類する。このとき、予測の確信度が高ければ、分類済みの訓練集合に追加し、以後の学習に用いている。また、予測の確信度を得るために、マージンに基づく分類器を用いている。

しかし、Levinらの手法は二つの主要な問題点をもっている。一つ目の問題点は、認識に用いられる複数の高レベル特徴が画素の階調値のみに基づいている点である。より汎用的な認識を行うためには、ほかにも形状に関する特徴など、多様な特徴を用いる必要がある。二つ目の問題点は、各分類器が相補的に学習していないという点である。多戦略学習では、各分類器が互いに助け合いながら学習を進めていくべきである。そのためには、相手の分類器にとって識別が困難な事例を集中的に学習する必要がある。

一つ目の問題を解決するためには、互いに大きく異なる特徴を用いて、多戦略学習を行う必要がある。多戦略学習に適している代表的な特徴として、画像の領域成分がある。Levinらの手法で用いられている画素の階調値は、領域成分に基づく特徴である。そのほかにも、色ヒストグラムなどがある[7]。しかし、これらは単純な特徴であるため、複雑な物体を記述するには不十分である。

また、物体の形状も、多戦略学習に適している特徴の一つである。例えば、Schieleらは物体の形状の確率的表現法を用いたlocal shape descriptorによる認識手法を提案している[8]。Belongieらは、物体の輪郭上の点を用いて形状記述子を生成し、分類に用いるshape context法を提案している[9]。

以上のことから、本論文では形状ベース及び領域ベースの認識手法を用いて多戦略学習を行う。形状ベースの認識手法では、物体の輪郭を断片的に抽出し[10]、群化に基づいて、意味のある断片の集合を発見し、物体を記述することにより、認識を行う。領域ベースの認識手法では、画像の周波数成分を解析するフーリエ変換[11]を拡張し、画像の回転などに対して頑健なGeneric Fourier Descriptor[12]と呼ばれる記述子を用いて認識を行う。

二つ目の問題点は、相手が誤分類した事例を集中的に学習することにより解決できる。具体的には、形状

(領域) ベースの分類器が誤分類した事例を、領域 (形状) ベースの分類器を訓練する際に重点的に学習すれば、互いに補い合いながら、相補的に学習することができる。

相補的な学習機構を確立するため、本手法は AdaBoost の学習機構に基づいている。しかし、AdaBoost には過学習を起しやすという問題点がある。このため、過学習を抑制する様々な手法が提案されている。例えば、Rätsch らは Soft Margin AdaBoost [13] を提案している。この手法では、重みを設定する際にパラメータを導入し、マージン最大化と過学習の抑制のトレードオフを図っているが、すべての事例の重みを決定する際に、このパラメータが用いられるので、過学習を引き起こす原因とならない事例の重みまで抑制してしまう可能性がある。また、Freund は、どの程度学習を繰り返すと過学習が発生するかを予測し、過学習が起きる前に学習を終了させる BrownBoost [14] を提案している。この手法では、過学習を防ぐことはできるが、訓練事例中に過学習の原因となる事例が多いときには、早い段階で学習を打ち切らなければならない。十分な学習ができない可能性もある。本手法では、過学習の原因となる事例の重みを選択的に抑制することにより、過学習の原因となる事例を発見する NadaBoost [15] を導入している。

また、物体認識課題は一般に多値問題であるのに対し、NadaBoost は 2 値分類問題を対象としているので、多値問題に適用するため拡張する必要がある。2 値分類アルゴリズムを用いて多値問題を解く一般的な手法として、多値問題を複数の 2 値問題に分解して解く AdaBoost.MO が提案されている [16]。しかし、この手法では、複数の 2 値分類問題の組合せを解く必要があり、本手法に導入すると計算コストの増大が問題となる。これに対し、AdaBoost アルゴリズムを直接多値問題に適用できるように拡張する AdaBoost.M1 が提案されている [17]。本論文でも AdaBoost.M1 と同様の立場をとり、NadaBoost アルゴリズムを多値問題に適用できるように拡張している。

3. 相補的学習

多様な特徴を用いた認識を適切に統合するための代表的な手法として、アンサンブル学習法 [5] がある。例えば、投票法は最も単純な統合法である。そのほかに、重み付けを行う投票法や AdaBoost [18] などがある。しかし、アンサンブル学習法では、各分類器が独立に

訓練され、他の分類器の学習結果を用いないので、相補的学習には不向きである。そこで、アンサンブル学習法を拡張し、相補的学習に適した学習法を構成する。

相補的学習では反復学習を行うため、AdaBoost [18] に基づいている。AdaBoost は、2 値分類問題に対する誤り率がただか $\frac{1}{2}$ の弱学習器からなる。反復により複数の弱学習器が生成され、それらによる重み付き投票により、最終仮説が生成される。重みは訓練事例に対する分類精度により決定される。つまり、大きい重みをもつ弱学習器は、最終仮説に多大な影響を与える。

相補的学習の最初のステップは、各学習器の欠点を見つけることである。そのために、各学習器が誤分類した事例を用いる。ある学習器が誤分類した事例は、その学習器にとって分類が困難であるため、他の分類器で重点的に学習する。逆に、その学習器では、他の学習器が誤分類した事例を重点的に学習する。こうして、複数の学習器が互いの欠点を補い合うようになっている。

3.1 NadaBoost

AdaBoost は、弱学習器が誤分類した事例に大きい重みを与え、それらを重点的に学習する。しかし、ノイズ事例や外れ値をもつ事例に過剰な重みを与えてしまい、結果として過学習を引き起こされることがある。そこで、過学習を防止するため、AdaBoost の重み付け規則を改善した NadaBoost を導入する [15]。

NadaBoost は訓練集合 $\{(x_1, y_1), \dots, (x_m, y_m)\}$ を入力にとる。ここで、 m は訓練事例数である。 x_i ($i = 1, \dots, m$) は事例空間 $\mathcal{X}$ (一般に、 d を任意の自然数とすると $x_i \in \mathbb{R}^d$) に属しており、 y_i ($i = 1, \dots, m$) はラベル空間 $\mathcal{Y}$ に属しているとする。2 値分類問題に対しては、 $\mathcal{Y} = \{-1, +1\}$ となる。また、物体認識においては、一つの事例が 1 枚の画像に対応している。更に、 x_i は画素の階調値のベクトル表現などの特徴ベクトルに相当している。したがって、その画像が認識対象の物体を含んでいれば “ $y_i = +1$ ”，含んでいなければ “ $y_i = -1$ ” となる。

AdaBoost では、ノイズ事例や外れ値をもつ事例である hard example が過学習の原因となる。物体認識では、ノイズのある画像や大部分が遮へいされている画像などがこれに相当している。そこで、NadaBoost では、この問題を解決するために重みのしきい値を決定し、しきい値を上回る事例を hard example とみなして、誤分類されても重みを減らしている。これによ

り, hard example に過剰な重みを与えないようにし, 過学習を防止している.

重みのしきい値は *HighWeight* と呼ばれ, 各ラウンドにおいて決定される. 第 t ラウンドにおける重みのしきい値 $HighWeight_t$ の候補は, 各事例の重みに基づいて決定される. 例えば, 訓練事例の重みが 0.1, 0.2, 0.3 のいずれかである場合, すなわち $\forall i D_t(i) \in \{0.1, 0.2, 0.3\}$ の場合, (1) $D_t(i) > 0.1$ となる事例が hard example であるとみなす, (2) $D_t(i) > 0.2$ となる事例が hard example であるとみなす, (3) すべての事例が hard example でないとみなす, という三つの可能性がある. したがって, しきい値の候補は, $HighWeight_t[1] = 0.1$, $HighWeight_t[2] = 0.2$, $HighWeight_t[3] = 0.3$ の三つとなる.

もし, 大部分の事例の重みが 0.1 であり, 0.2 以上の重みをもつ事例が少ない場合は, 重みが 0.1 を超える事例が hard example である可能性が高いので, $HighWeight_t[1]$ をしきい値として採用するのが妥当である. 一方, 重みが 0.3 の事例は少ないが, 0.2 以下の重みをもつ事例が多い場合には, 重みが 0.3 の事例が hard example である可能性が高いので, $HighWeight_t[2]$ を採用すべきである. このように, 与えられる訓練集合によって最適な値が異なるため, しきい値を決定するのは難しい.

そこで, ビームサーチを用いて, これらのしきい値候補に対して探索を行う. このとき, 訓練事例に対する誤分類率の理論値が最小となり, かつ AdaBoost の誤分類率を下回るように定めている. *HighWeight* を用いた NadaBoost の重み更新規則は以下のように表される.

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i) I_{t,i})}{Z_t} \quad (1)$$

ここで, h_t , $D_t(i)$ は, それぞれ第 t ラウンドでの弱仮説及び i 番目の事例に与えられる重みを表し, α_t は弱仮説の分類精度に基づく重みを表す. また,

$$I_{t,i} = \begin{cases} -1 & \text{if } y_i h_t(x_i) < 0 \\ & \wedge D_t(i) > HighWeight_t \\ +1 & \text{otherwise} \end{cases}$$

とする. 重みが *HighWeight* を超える事例は hard example とみなされ, 正しく分類された事例と同様に重みが減らされるので, 過学習を防ぐことになる.

3.2 多値問題への拡張

NadaBoost は, 元来, 2 値問題を対象としているが, 物体認識は一般に多値問題なので, 拡張する必要がある. はじめに, ラベル空間を $\mathcal{Y} = \{1, \dots, C\}$ (C はクラス数) と定める. つまり, C 種類の物体を識別する問題を考える. 多値問題では, 弱仮説 $h_t(x)$ は, 与えられた画像 x がどの物体であるかを予測することになる. したがって, 式 (1) を用いて, NadaBoost の重み付け規則を以下のように書き直している.

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t b_{t,i} \hat{I}_{t,i})}{Z_t} \quad (2)$$

ただし, $b_{t,i} = \langle y_i = h_t(x_i) \rangle$ とおき,

$$\hat{I}_{t,i} = \begin{cases} -1 & \text{if } b_{t,i} = -1 \\ & \wedge D_t(i) > HighWeight_t \\ +1 & \text{otherwise} \end{cases}$$

とする. また, 任意の述語 π について, π が真のとき $\langle \pi \rangle = +1$, π が偽のとき $\langle \pi \rangle = -1$ とする. 以上より, 最終仮説 H は次の式で与えられる.

$$H(x) = \operatorname{argmax}_c \sum_{t=1}^T \alpha_t \langle h_t(x) = c \rangle. \quad (3)$$

多値問題に拡張した NadaBoost は, 重み付け規則と最終仮説を除いて, 通常の NadaBoost と同じである.

3.3 NadaBoost を用いた相補的学習

本節では, 多戦略学習を行うために, NadaBoost アルゴリズムを拡張し, 相補的学習法を提案する. アルゴリズムの記述を簡潔にするため, ここでは二つの物体認識法 A と B (これらを base learner と呼ぶ) を用いた相補的学習について述べる.

各ラウンドで, 各訓練事例の重みを更新する際, A と B は互いに相手の学習結果を用いて相補的に学習を行う. 具体的には, 第 $(t+1)$ ラウンドで B の弱仮説 h_{t+1}^B を生成するとき, 第 t ラウンドで A の弱仮説 h_t^A が誤分類した事例の重み D_{t+1}^B が増やされる. A の弱仮説 h_{t+1}^A も同様にして生成される. ここで, 相補的学習のための重み付け規則を以下に示す.

$$D_{t+1}^A(i) = \frac{D_t^B(i) \exp(-\alpha_t^B b_{t,i}^B \hat{I}_{t,i}^B)}{Z_t^B} \quad (4)$$

$$D_{t+1}^B(i) = \frac{D_t^A(i) \exp(-\alpha_t^A b_{t,i}^A \hat{I}_{t,i}^A)}{Z_t^A} \quad (5)$$

$$\alpha_t^X = \frac{1}{2} \ln \left(\frac{1 + r_t^X}{1 - r_t^X} \right) \quad (6)$$

ただし, $X \in \{A, B\}$, $b_{t,i}^X = \langle y_i = h_t^X(x_i) \rangle$ とし,

$$r_t^X = \sum_{i=1}^m D_t^X(i) b_{t,i}^X.$$

である。ここで, D^X , α^X , h^X , $\hat{f}^X$, Z^X はそれぞれ base learner X の D , α , h , $\hat{f}$, Z に相当する。式 (4), (5) より, D_{t+1}^A が B の, D_{t+1}^B が A の学習結果に依存していることが分かる。この重み付け規則により, A , B は相手が誤分類した事例を集中的に学習できる。

以上の重み付け規則により計算された重みを用いて, 第 t ラウンドにおける, 訓練事例に対する base learner X の仮説の分類誤り上限 $estimated_error_t^X$ を以下のように求めることができる。

$$estimated_error_t^X = \left(\prod_{k=1}^t \sqrt{1 - (r_k^X)^2} \right) \cdot \left\{ \sum_{i=1}^m D_{t+1}^X(i) \cdot \exp \left(2 \sum_{k: \hat{f}_{k,i}^X = -1} \alpha_k^X \right) \right\} \quad (7)$$

また, AdaBoost の訓練誤りの上限値 $errorbound_ada^X$ も, 各ラウンドで生成された重みを用いて, 以下のように求めることができる。

$$errorbound_ada^X = \prod_{t=1}^T \sqrt{1 - (r_t^X)^2}. \quad (8)$$

重みのしきい値 $HighWeight$ を決定するときに, $estimated_error$ と $errorbound_ada$ の値を比較することにより, 仮説の訓練誤りが AdaBoost の訓練誤りを下回るように $HighWeight$ を設定できる。

以上より, 最終仮説は以下の式で与えられる。

$$H(x) = \argmax_c \left\{ s^A(c) \sum_{t=1}^T \alpha_t^A \langle h_t^A(x) = c \rangle + s^B(c) \sum_{t=1}^T \alpha_t^B \langle h_t^B(x) = c \rangle \right\}. \quad (9)$$

上式において, s^A , s^B はそれぞれ A , B のクラス分離度を表している。分離度はそのクラスに属する事例がどれだけ他のクラスと混同されにくいかを示す指標である。クラス分離度は A と B に対する混同行列 $[n_{i,j}^A]$ 及び $[n_{i,j}^B]$ を用いて次のように計算される。

$$s^A(c) = k \frac{\nu_c^A}{\sum_{i=1}^C \nu_i^A}, \quad s^B(c) = k \frac{\nu_c^B}{\sum_{i=1}^C \nu_i^B} \quad (10)$$

k は $s^A(c) + s^B(c) = 1$ となるように定められ,

$$\nu_c^A = \frac{n_{c,c}^A}{\sum_{i=1}^C n_{c,i}^A}, \quad \nu_c^B = \frac{n_{c,c}^B}{\sum_{i=1}^C n_{c,i}^B}.$$

である。また, 混同行列の要素 $n_{i,j}$ は, クラス i に属しており, クラス j に分類された訓練事例の数を表す。クラス分離度が高ければ分類精度も高いため, クラス分離度は, 最終仮説を決定する際に, より適切な base learner の予測を重視するために用いられる。

最後に, 相補的な視覚的学習アルゴリズムを示す。ここで, 第 t ラウンドにおけるしきい値候補の数を d_t と表記する。また, 各 base learner $X \in \{A, B\}$ に対する事例の重み $D_t^X(i)$, 弱仮説 h_t^X , α_t^X , $estimated_error_t^X$ の四つ組を「状態」と定義し, 第 t ラウンドでの j 番目のしきい値候補 $HighWeight_t[j]$ に対応する状態を $S_t[j]$ ($j = 1, \dots, d_t$) と表記する。更に, 探索対象となる状態を保持しておくリストを S_list , 最終状態 (T ラウンド終了後の状態) を保持しておくリストを $Goal_list$ と呼ぶ。

1. AdaBoost を実行し, 訓練誤りの上限値 $errorbound_ada$ の値を得る。
また, $D_1^A(i) = D_1^B(i) = \frac{1}{m}$, $t = 1$, $d_1 = 1$, $HighWeight_1[1] = \frac{1}{m}$ とし, S_list 及び $Goal_list$ は空とする。
2. A , B の弱仮説 h_t^A , h_t^B を生成する。
3. $HighWeight_t[j]$ ($j = 1, \dots, d_t$) を第 t ラウンドでの j 番目のしきい値候補とする。
4. 各 $HighWeight_t[j]$ に対して, α_t^A , α_t^B , D_{t+1}^A , D_{t+1}^B が, 式 (4), (5), (6) で表される修正された重み付け規則に従って計算される。そして, 仮説の分類誤りの上限 $estimated_error_t^A$, $estimated_error_t^B$ が式 (7) を用いて計算される。これらの値と弱仮説は状態 $S_t[j]$ が保持する。
5. 各状態 $S_t[j]$ を状態リスト S_list に追加する。ただし, S_list は最大 $beam_width$ 個までしか状態を保持しないので, 保持すべき状態数が $beam_width$ 個を上回る場合は, $estimated_error$ の大きいものから順に状態を削除する。
6. S_list が空であれば 8 へ進む。
7. 最小の $estimated_error$ をもつ $S_k[j]$ について,

$k < T$ の場合は, $t \leftarrow k + 1$ とし, その状態を用いて 2 へ進む. $k = T$ の場合は, $estimated_error^A < errorbound_ada^A$ かつ $estimated_error^B < errorbound_ada^B$ であれば, $S_k[j]$ を *Goal_list* に追加して 6 へ進む.

8. *Goal_list* が空でなければ, *Goal_list* に含まれている状態のうち, $estimated_error^A + estimated_error^B$ が最小になる状態が選択され, 最終仮説がすべての弱仮説の予測の和として生成される. このとき, クラス分離度を用いて重み付けが行われる. *Goal_list* が空であれば, はじめに実行した AdaBoost の最終仮説を出力する.

4. 視覚的学習法

4.1 形状ベースの視覚的学習

輪郭などの物体の形状は, 物体の認識に非常に有効な要素である. 実際, 人の視覚も物体の形状を認識に用いていることが知られており, ゲシュタルト心理学の創始者 Wertheimer は次のように述べている [4]. 人がある物体を見たとき, 最初にその輪郭や領域の断片が知覚される. しかし, これらの断片は, その物体に関する十分な情報をもたないため, 単独では認識に用いることができない. そこで, 次に人は複数の断片を無意識にはあるが注意深く選択し, 全体として意味のある集合を形成する. これにより, 物体が意味のあるまとまりとして知覚される.

このような処理は「群化」と呼ばれており, 人のもつ主要な視覚的能力の一つである. 形状ベースの学習として, 群化の概念に基づいた認識手法を構成する. 具体的には, まず物体の輪郭を複数の直線からなる集合を用いて記述する. この直線を輪郭片と呼ぶ. 輪郭片は単純な直線で表されるので, ごく少数の輪郭片で複雑な物体を表現することはできない. しかし, 複数の輪郭片間の位置関係を考えると, 物体の特徴をとらえた構造を発見できる. 例えば, 直交や平行など規則的に並んでいる輪郭片は, 人工物に多く含まれる. 逆に, 自然物には曲線的に並んでいたり不規則に並んでいる輪郭片が多い. このような輪郭片の位置的關係は, 両者を識別するのに有効である. したがって, 複数の輪郭片から意味のある構造 (これをパターンと呼ぶ) を発見し, その中でも認識対象の物体に頻繁に含まれるものを, 有用なパターンとして認識に利用する.

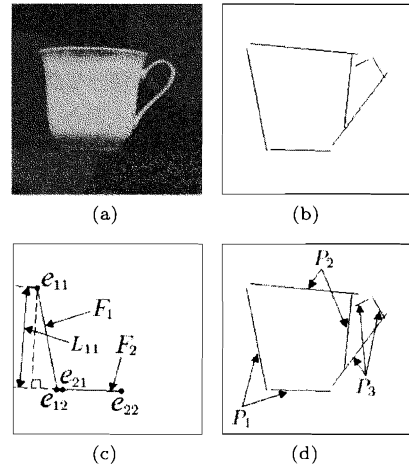


図 1 パターン抽出の例

Fig. 1 An example of the pattern extraction process.

学習プロセスにおいて, はじめに与えられた画像から輪郭片を抽出し, 得られた輪郭片の集合の中から, 意味のある集合, すなわちパターンを発見する. 図 1 に簡単な例を示す.

図 1 (a) は与えられた画像を示しており, (b) が (a) から抽出された輪郭片の集合を示している^(注1). 輪郭片の抽出には, Nelson による stick growing 法を用いている [19]. stick growing 法では, 物体のエッジの端点及びこう配を求め, 物体の輪郭を複数の直線の集合として抽出する. 本手法では, 入力画像をグレースケールに変換した画像を stick growing 法による輪郭抽出アルゴリズムの入力とし, 得られた直線の集合を輪郭片の集合として用いている.

次に, 群化の概念に従い, 互いに近接している輪郭片をグループ化してパターンを生成する. 二つの輪郭片が互いに近接しているかを判定するための条件は以下のように定める.

$$\min\{L_{11}, L_{12}, L_{21}, L_{22}\} < d$$

ここで, d はしきい値である. 二つの輪郭片を F_1, F_2 と記述し, 各輪郭片 F_i の端点を e_{i1}, e_{i2} ($i = 1, 2$) とする. L_{ij} は, 一方の輪郭片の端点 e_{ij} から他方の輪郭片 (の延長線上) に引いた垂線の長さを表す. 例えば, L_{11} は輪郭片 F_1 の端点 e_{11} から他方の輪郭片 F_2 の延長線上に引いた垂線の長さを表す. (c) は互いに近接している輪郭片の組を表している. 二つの輪郭片

(注1): 図を見やすくするため, 図 1 (b) には, 背景などから抽出された有用でない輪郭片は示していない.

F_1, F_2 が上記の条件を満たせば、これらは隣接しており、意味のある組合せとみなされる。(d) は (b) から発見された三つのパターン P_1, P_2, P_3 を示している。このようにして群化を行い、パターンを得る。

ある特定のクラスに属する物体に頻繁に含まれるパターンは、その物体の特徴的な部分をとらえていると考えられるため、有用なパターンである。これらを頻出パターンと呼び、以下の条件を満たすものと定める。

$$\frac{n_c}{N_c} > \rho \quad (11)$$

ここで、 n_c はクラス c に属する事例（物体）の中で、このパターンを含むものの数を表す。 N_c はクラス c に属する事例の総数を表す。 ρ は頻度のしきい値である。

頻出パターンは、そのクラスだけでなく、ほかのクラスに属する物体にも多く含まれていると、識別に利用できない。つまり、有用なパターンとは、ある特定のクラスに属する物体にのみ頻出し、それ以外の物体にはほとんど含まれないパターンである。そこで、頻出パターンの有用性を評価するため、クラス c の頻出パターン p の有用度 $U_c(p)$ を次のように定める。

$$U_c(p) = \frac{n_c}{\sum_{i=1}^C n_i}$$

ここで、 C はクラス数を表し、 n_i はクラス i に属しており、かつ p を含む事例数を表す。クラス c に属するすべての事例が p を含んでおり、そのほかのクラスに属するすべての事例が p を含んでいない場合、有用度は最大で 1 となる。この値は、二つの物体間の類似度を計算する際に重みとして用いられる。

未知の事例は、得られた頻出パターンの有用度に基づいて分類される。未知の事例が与えられたとき、式 (11) を用いて頻出パターンの集合 $\{p_i\}$ ($i = 1, \dots, m$, m は頻出パターン数) が抽出される。各パターン p_i は、各クラスの頻出パターン q_i^c ($i = 1, \dots, m_c$, m_c はクラス c の頻出パターン数) と比較される。 p_i, q_i 間の類似度 $\sigma(p_i, q_i)$ は、各頻出パターンについて計算される。そして、類似度を用いて確信度 S_c が求められる。 S_c は、未知の事例がクラス c に属する確率に相当し、次のように類似度の和を用いて計算される。

$$S_c = \sum_{i=1}^M \sigma(p_i, q_i) \quad (12)$$

$$\sigma(p_i, q_i) = \begin{cases} U_c(q_i) & \text{if } p \text{ is similar to } q \\ 0 & \text{otherwise} \end{cases}$$

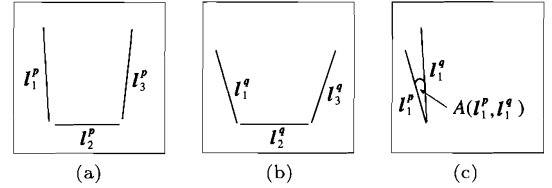


図 2 パターン p と q の例
Fig. 2 An example of patterns p and q .

$$M = \min\{m, m_c\}$$

各クラスに対する確信度は、事例のクラスの予測に用いる。確信度を計算する際に、式 (12) を最大化するように、各 i に対して p_i に対応する q_i を定める。

パターン p がパターン q に類似しているかどうかを決定するために、 p, q に含まれる各輪郭片 l_i^p, l_i^q について、以下の三つの条件を定める。

条件 1 $n_p = n_q = n$.

条件 2 $\frac{1}{r} < |l_i^p|/|l_i^q| < r$ ($i = 1, \dots, n$).

条件 3 $A(l_i^p, l_i^q) < \theta$ ($i = 1, \dots, n$).

ここで、 n_p, n_q は、それぞれパターン p, q に含まれる輪郭片の数を表す。また、 $A(x, y)$ は x と y のなす角を表し、 r, θ はしきい値である。 p と q に含まれるすべての輪郭片が上記の条件を満たしていれば、両者は類似しているとみなされる。図 2 に例を示す。

図 2 の (a), (b) はそれぞれパターン p, q を表している。両パターンは三つの輪郭片で構成されているので、条件 1 は満たされている。(c) は $i = 1$ に対応する輪郭片の組を表している。輪郭片の長さの比 $|l_1^p|/|l_1^q|$ 及び両者のなす角 $A(l_1^p, l_1^q)$ がそれぞれ条件 2, 3 を満たしており、かつ $i = 2, 3$ についても各条件が満たされていれば、 p と q は類似していると判断される。以上の類似条件を用いて、各クラスに対する類似度を求め、次の式で示されるクラス c^* に事例进行分类する。

$$c^* = \underset{c}{\operatorname{argmax}} S_c. \quad (13)$$

すなわち、確信度が最も高いクラスに分類する。

4.2 領域ベースの視覚的学習

物体の色やテクスチャを表す領域成分も、物体の識別に有効な特徴である。物体は、画素の階調値に関して一様な分布をもついくつかの領域に分割される。本手法では、すべての輪郭片を包含する最小の円形領域内の画素の分布を用いる。図 3 にこの領域の例を示す。

図 3 (a) 内の円で囲まれた領域が最小領域である。(b) は輪郭片に対応する最小領域を示している。最小

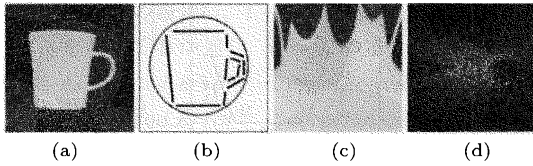


図3 最小領域と GFD の例

Fig. 3 An example of the minimum region and GFD.

領域内に含まれる画素の階調値により、領域成分が表される。しかし、画素数が多いため、すべての画素を用いると、特徴の次元数が高くなり、計算量が非常に増大する。そこで、極フーリエ変換を用いて特徴ベクトルを生成する。これにより、物体の特徴を保持しつつ、特徴ベクトルの次元数を抑えることができる。そのために、Generic Fourier Descriptor (GFD) [12] と呼ばれる特徴記述子を導入する。GFD は、画像の極座標表現に極フーリエ変換を適用し、フーリエ係数を適切に正規化すると得られる特徴ベクトルである。

GFD を求めるためには、まず図 3(c) に示すように、画像の極座標表現を行う。画像 (a) の極座標表現 (c) は、(a) の最小領域の中心を原点として、水平軸方向に角度、垂直軸方向に中心からの距離をとると得られる。次に、極座標表現された画像に、以下の二次元極フーリエ変換 F を適用して極フーリエ係数を得る。

$$F(\rho, \phi) = \sum_{r=0}^{R-1} \sum_{i=0}^{T-1} \left\{ f(r, \theta_i) \cdot \exp \left[j2\pi \left(\frac{r}{R}\rho + \frac{2\pi i}{T}\phi \right) \right] \right\} \quad (14)$$

ここで、 R 及び T はそれぞれ径（垂直）方向及び角（水平）方向の解像度を表す。また、 $\theta_i = \frac{2\pi i}{T}$ ($0 \leq i < T$) とする。図 3(d) は式 (14) に従って計算された極フーリエ係数を表している。以下のように特徴ベクトル GFD を計算することにより、画像の回転に対して頑健な特徴が得られる。

$$GFD = \left\{ \frac{|F(0,0)|}{area}, \frac{|F(0,1)|}{|F(0,0)|}, \dots, \frac{|F(0,n)|}{|F(0,0)|}, \dots, \frac{|F(m,0)|}{|F(0,0)|}, \dots, \frac{|F(m,n)|}{|F(0,0)|} \right\} \quad (15)$$

ここで、 $area$ は画像の面積を表し、 m, n はそれぞれ径、角周波数の最大値を表す。以上のようにして、最小領域内の画素の分布に基づく特徴ベクトルが得られる。最終的に、GFD 特徴ベクトル間の距離を用いて、画像間の類似度が決定される。

未知の事例は、類似度に従って分類される。まず、

最小領域とその極座標表現を求める。次に、式 (14) より極フーリエ係数を得る。最後に、式 (15) を用いて GFD 特徴ベクトルを計算する。未知の事例 x と訓練事例 t の間の類似度 $S(x, t)$ は、GFD 特徴ベクトル間のユークリッド距離を用いて次のように定める。

$$S(x, t) = \left(\sum_{i=1}^d \{GFD_i(x) - GFD_i(t)\}^2 \right)^{-\frac{1}{2}} \quad (16)$$

ここで、 d は GFD 特徴ベクトルの次元数を表す。また、 $GFD_i(x)$, $GFD_i(t)$ はそれぞれ x と t の GFD 特徴ベクトルの i 番目の値を表す。未知の事例 x は以下の式で示されるクラス c^* に分類される。

$$c^* = \operatorname{argmax}_c \frac{\sum_{i=1}^{n_c} S(x, t_i^c)}{n_c} \quad (17)$$

ここで、 t_i^c はクラス c に属する i 番目の訓練事例を表し、 n_c はクラス c に属する訓練事例数を表す。事例 x は類似度の平均値が最大のクラスに分類される。

5. 実 験

提案手法の性能を評価するため、実世界の画像を用いて認識実験を行う。画像は ETH-80 Image Set database [20] より取得しており、8 種類の物体 apple, car, cow, cup, dog, horse, pear, tomato からなる。画像数は各クラス 410 枚であり、それぞれ 10 個の異なる物体を 41 の視点から撮影した画像を含んでいる。

5.1 既存の認識手法との比較

まず、提案手法の有用性を検証するために、既存の物体認識手法との認識精度の比較を行う。本実験では、訓練集合として 656 枚（各クラス 82 枚）の画像を使用し、残りの 2624 枚（各クラス 328 枚）の画像をテスト集合として使用している。また、ラウンド数は 100、ビーム幅 (*beam_width*) は 10, 30, 50 としている。表 1 に、5 分割の交差検定を用いた実験結果を示す。

表 1 より、ビーム幅を大きくすると、全体の分類精度が増加することが分かる。これは、相補的学習を行うと、通常の NadaBoost より探索空間が広くなり、ビーム幅が小さいと十分な探索ができないためである。また、動物 (cow, dog, horse) の分類精度が比較的低い。これは、同じ種類の動物でも、個体によって色や模様が大きく異なり、領域ベースの手法では有効な特徴の発見が困難なためである。更に、動物の形状は、

表 1 分類精度 (単位は%)
Table 1 The recognition accuracy (in %).

classes	beam_width		
	10	30	50
apple	86.23	89.93	92.04
car	84.59	88.51	88.40
cow	59.42	63.31	65.65
cup	76.20	82.80	84.85
dog	60.07	66.81	70.16
horse	64.27	68.73	71.54
pear	80.80	83.16	84.58
tomato	70.18	73.65	75.44
total	72.72	77.11	79.08

姿勢や視点によって異なるため、形状ベースの手法でも正確な認識が難しいことも原因である。したがって、形状・領域ベースのどちらを用いても、動物を識別するための有用な特徴をとらえることはできていない。

次に、提案手法と既存手法の間で、分類精度の比較を行う。まず、次の二つの形状ベースの認識手法との比較を行う。一つは、Schiele らによる multidimensional receptive histogram [8] であり、統計的表現を用いて物体の局所的形状を記述し、認識に用いる手法である。もう一つは、Belongie らによる shape context [9] であり、輪郭線上に均等に配置した点を用いて物体の形状を記述し、マッチングによる認識を行う手法である (注2)。

次に、次の二つの領域ベースの認識手法との比較を行う。一つは、Swain らによる color indexing [7] であり、物体を構成するすべての画素の RGB ヒストグラムを用いて認識を行う手法である。もう一つは、Grauman らによる local invariant features [21] であり、こう配に基づく記述子により物体の局所的不変特徴を発見し、認識に用いる手法である。

最後に、機械学習を用いている次の二つの認識手法との比較を行う。一つは、Marée らによる random subwindows [2] であり、画像内からランダムに抽出された、複数のスケールのサブウィンドウを用いて、画像を分類する決定木のアンサンブルを構築する手法である。もう一つは、Tu による probabilistic boosting-tree [3] であり、決定木と boosting を組み合わせた、確率的 boosting 決定木により識別モデルを構成する手法である。各手法による分類精度を表 2 に示す。

表 2 より、提案手法の認識精度は、既存手法と比べてそれほど高いとはいえないが、最も高い認識精度を達成している shape context 法と比べても、認識精度の差は小さい。また、提案手法の base learner の認識

表 2 各手法の分類精度
Table 2 The recognition accuracy for each method.

	accuracy (%)
Swain (1991)	64.85
Marée (2005)	74.51
Tu (2005)	76
proposed method	79.08
Schiele (2000)	79.79
Grauman (2005)	81
Belongie (2001)	81.06

精度は、いずれも単体では Marée らの手法よりも若干低い程度なので、これらを統合することによる認識性能の向上は大きい。したがって、この結果は多戦略学習の有効性を示している。

5.2 NadaBoost の有効性

NadaBoost が AdaBoost よりも hard example に対して頑健であることを示すために、訓練事例に人工的に hard example を混入したデータセットを用いて、比較実験を行った。本実験に用いた訓練集合は、全訓練事例のうち、ランダムに 10% の事例を選択し、そのラベルを正しくないものに書き換え、hard example としている。また、テスト集合は、前述の実験と同一のものを用いている。

図 4 に、5 分割の交差検定を用いた際の、AdaBoost 及びビーム幅を 10, 30, 50 とした NadaBoost の認識精度を示す。

100 ラウンド終了時における認識精度は、すべてのビーム幅の値について、NadaBoost が AdaBoost を上回っている。また、ビーム幅を 30, 50 としたときの NadaBoost は、有意水準を 5% とした *t* 検定において有意な差が生じている。以上の結果より、訓練事例に hard example が含まれる場合でも、NadaBoost は AdaBoost よりも高い認識精度を達成できることが確認できた。

更に、AdaBoost では、60 ラウンドを過ぎたあたりから、ほとんど認識精度の改善が見られなくなっているのに対し、NadaBoost では、100 ラウンドに至る

(注2)：[20]において、Belongie らの shape context 法の認識精度は 86.40% であると報告されている。しかし、これはデータセットに含まれる事例の 98% 以上を訓練事例として用いた際の認識精度である。Shape context 法では、訓練事例とのマッチングにより認識を行うため、[9]にも示されているように、認識精度が訓練事例数に比例する。したがって、この値は shape context 法に有利な実験の結果である。一方、本手法の base learner を訓練するときに用いている訓練事例数は、データセットに含まれる事例の 20% であり、[20]の実験で用いられている数に比べて非常に少ない。そこで、本論文では、データセット中の 20% の事例を訓練事例とした際の認識精度を示している。

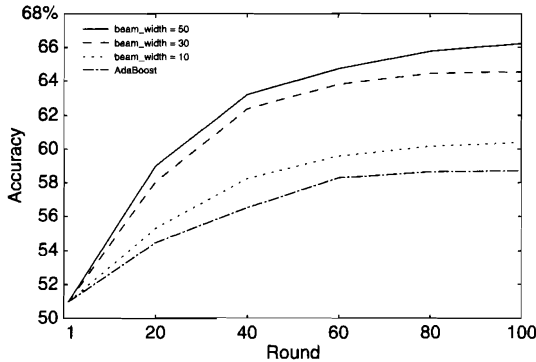


図 4 ノイズを含む訓練事例を用いた際の認識精度

Fig. 4 The recognition accuracy using noisy training set.

表 3 提案手法と投票法の認識精度 (%)

Table 3 The recognition accuracy (in %) of proposed method and voting method.

クラス	相補的学習	投票法	t
apple	89.93	83.74	*
car	88.51	84.47	*
cow	63.31	63.29	-
cup	82.80	79.75	*
dog	66.81	63.75	*
horse	68.73	64.87	*
pear	83.16	79.00	*
tomato	73.65	72.84	-
total	77.11	73.96	*

まで、ラウンド数が増加するにつれて認識精度が向上している。したがって、AdaBoost がノイズ事例による過学習の影響を受け、認識精度が抑制されてしまうのに対し、NadaBoost はノイズ事例を適切に判断し、その重みを減らすことができている、過学習の影響を抑えられていることが確認できた。

5.3 相補的学習の有効性

相補的学習の有効性を検証するため、提案手法と単純な投票法との比較を行う。投票法では、形状・領域ベースの認識手法は別々に訓練され、重み付けなしの投票法により、各手法の学習結果が統合される。すなわち、投票法での重み付け規則は次のように表される。

$$D_{t+1}^X(i) = \frac{D_t^X(i) \exp(-\alpha_t^X b_{t,i}^X \hat{f}_{t,i}^X)}{Z_t^X}.$$

ここで、 X は形状または領域ベースの手法にあたる。

実験結果を表 3 に示す。 t 列の * は、提案手法と投票法の認識精度の間に、有意水準 5% での t 検定において有意な差があることを示している。また、ビーム幅はともに $beam_width = 30$ としている。

すべての物体に対する認識精度及び全体の認識精度

について、提案手法は投票法を上回っている。更に、そのうちの多くに有意な差が見られる。ある物体が、形状 [領域成分] に特徴的な要素をもたない場合、形状 [領域] ベースの認識手法では、誤識別する可能性が高い。しかし、そのような場合でも、投票法ではすべての学習器の予測を等しく扱うので、形状 [領域] ベースの学習器の予測が悪影響を及ぼす可能性が高い。これに対し、相補的学習では、適切な学習器の予測を重視できるため、この悪影響を低減できる。この結果は、相補的学習の長所を反映しているといえる。

5.4 多数の base learner による相補的学習

3. では、二つの base learner による相補的学習アルゴリズムについて述べたが、本アルゴリズムは、多数の base learner を用いた一般的な多戦略学習に拡張できる。 n 個の base learner を用いた相補的学習を行う場合、NadaBoost における l 番目の base learner の重み付け規則は、以下のように表される。

$$D_t^l(i) = \frac{D_t^{f(t)}(i) \exp(-\alpha_t^{f(t)} b_{t,i}^{f(t)} \hat{f}_{t,i}^{f(t)})}{Z_t^{f(t)}}$$

$$f(t) = \begin{cases} M+1 & \text{if } M \leq l-2 \\ M+2 & \text{if } M \geq l-1 \end{cases}$$

ここで、 $M = t \bmod (n-1)$ とする。また、“ $x \bmod y$ ” は、 x/y の剰余を表す。すなわち、各 base learner は、自身以外の base learner の学習結果を順次用いて重みを更新する。例えば、base learner が四つの場合、2 番目の base learner について、第 1 ラウンドでは $M = 1 \geq l-1 = 1$ となるので、 $f(1) = M+2 = 3$ となる。同様に、第 2 ラウンドでは $M = 2 \geq 1$ となるので、 $f(2) = 4$ となる。第 3 ラウンドでは、 $M = 0 \leq l-2 = 0$ となるので、 $f(3) = M+1 = 1$ となる。したがって、第 1, 2, 3 ラウンドでそれぞれ 3, 4, 1 番目の base learner の学習結果に基づいて重みを更新する。 $f(t)$ の値は、 $(n-1) = 3$ ラウンド周期で繰り返されるので、以降のラウンドでは、3, 4, 1 番目の base learner の学習結果を、この順序で繰り返し用いて学習していく。以上の重み付け規則より、最終仮説 $H(x)$ は次の式で与えられる。

$$H(x) = \operatorname{argmax}_c \left\{ \sum_{l=1}^n \left(s^l(c) \cdot \sum_{t=1}^T \alpha_t^l \langle h_t^l(x) = c \rangle \right) \right\}. \quad (18)$$

表 4 三つの base learner による相補的学習の分類精度
Table 4 The recognition accuracy (in %) of the complementary learning with 3 base learners.

classes	2 learners	3 learners	t
apple	89.93	89.62	-
car	88.51	90.15	-
cow	63.31	66.21	*
cup	82.80	86.23	*
dog	66.81	67.81	-
horse	68.73	70.78	*
pear	83.16	85.89	*
tomato	73.65	78.09	*
total	77.11	79.35	*

式 (18) における l 番目の base learner のクラス分離度 s^l の定義は、式 (10) より導出可能なため省略する。

本節では、新たに特徴木に基づく認識手法 [1] を追加し、三つの base learner を用いた相補的学習の有効性を検証する。この認識手法では、画素の分布などの低レベル特徴を統合し、特徴木と呼ばれる木構造を生成する。特徴木は一種の決定木であり、各節点でそれぞれ異なる低レベル特徴を用いて、画像を分類する。

両者の比較結果を表 4 に示す。表中の “2 learners” と “3 learners” の列は、それぞれ二つ及び三つの base learner による相補的学習の結果を示す。t 列の * は、両者の認識精度の間に、有意水準 5% での t 検定で有意な差があることを示している。また、ビーム幅はともに $beam_width = 30$ としている。

全体の分類精度は base learner の追加により改善され、有意な差も生じている。また、半数以上のクラスに対する分類精度に有意な差が見られる。特に cow, horse, tomato については、base learner の増加による相補的学習の効果がよく現れている。この結果は、多様な学習手法の統合が分類性能の向上につながるという、アンサンブル学習の原理に従っている。

6. む す び

本論文では、多様な高レベル特徴を用いた多戦略学習に基づく視覚的学習法を提案した。更に、効果的な多戦略学習のため、相補的学習法を提案した。本手法では、複数の base learner が互いに欠点を補い合いながら学習できるので、多様な base learner を導入することにより、学習課題に依存しない、汎用的な認識機構を構成することができる。また、base learner として、二種類の高レベル特徴に基づく物体認識手法を提案した。一つは、形状に基づいており、群化の概念を用いて、物体の輪郭の断片をグループ化し、特徴を生

成する手法である。もう一つは、領域成分に基づいており、GFD を用いて、画素の分布を簡潔に表せる特徴を生成する手法である。

提案手法の性能を検証するため、実世界の画像を用いた実験を行った。Base learner が二つのときは、提案手法は既存手法と比較してそれほど高い認識精度を達成しているとはいえないが、base learner 単体での認識精度に比べると、大きく改善することができた。更に、base learner を三つ用いた相補的学習の有用性を検証した。実験結果より、より多様な認識手法を組み合わせると、分類精度の改善につながる事が確認できた。

しかし、提案手法には、クラス内に多様な形状やテクスチャをもつ物体を十分識別できないという問題点が残されている。この問題点は、特に動物の分類において顕著に見られる。そこで、適切に人の視覚モデルを定式化することにより、base learner となる物体認識手法を改善する必要がある。

また、NadaBoost を多値問題に拡張する際に、AdaBoost.M1 アルゴリズムに基づいた拡張法を用いているため、個々の base learner の認識精度が低い場合に、学習誤り率を減らすことが困難になるという問題がある。この問題を解決するために、AdaBoost.M2 アルゴリズムなどに基づいて、base learner の認識性能が低い場合でも、効果的な学習を行えるように NadaBoost アルゴリズムを改善すべきである。更に、NadaBoost は最適な重みのしきい値を探索するため、AdaBoost に比べて計算量が大きいという問題がある。そこで、多数の base learner を用いて効率的に学習を行うために、計算効率の改善を行うことも今後の課題である。

文 献

- [1] H. Nomiya and K. Uehara, “Feature construction and feature integration in visual learning,” Proc. ECML2005 Workshop on Sub-symbolic Paradigm for Learning in Structured Domains, pp.86–95, 2005.
- [2] R. Marée, P. Geurts, J. Piater, and L. Wehenkel, “Random subwindows for robust image classification,” Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, pp.34–40, 2005.
- [3] Z. Tu, “Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering,” Proc. 10th IEEE International Conference on Computer Vision, vol.2, pp.1589–1596, 2005.
- [4] D. Marr, Vision, W.H. Freeman and Company, 1982.
- [5] T.G. Dietterich, “Ensemble methods in machine

- learning," Proc. 1st International Workshop on Multiple Classifier Systems, pp.1-15, 2000.
- [6] A. Levin, P. Viola, and Y. Freund, "Unsupervised improvement of visual detectors using co-training," Proc. 9th IEEE International Conference on Computer Vision, pp.626-633, 2003.
- [7] M.J. Swain and D.H. Ballard, "Color indexing," Int. J. Comput. Vis., vol.7, no.1, pp.11-32, 1991.
- [8] B. Schiele and J.L. Crowley, "Recognition without correspondence using multidimensional receptive field histograms," Int. J. Comput. Vis., vol.36, no.1, pp.31-50, 2000.
- [9] S. Belongie, J. Malik, and J. Puzicha, "Matching shapes," Proc. 8th IEEE International Conference on Computer Vision, pp.454-463, 2001.
- [10] A. Selinger and R.C. Nelson, "Improving appearance-based object recognition in cluttered backgrounds," Proc. 15th International Conference on Pattern Recognition, vol.1, pp.46-50, 2000.
- [11] D. Zhang and G. Lu, "A comparative study of Fourier descriptors for shape representation and retrieval," Proc. 5th Asian Conference of Computer Vision, pp.646-651, 2002.
- [12] D. Zhang and G. Lu, "Enhanced generic Fourier descriptors for object-based image retrieval," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp.3668-3671, 2002.
- [13] G. Rätsch, T. Onoda, and K.-R. Müller, "Soft margins for adaboost," Mach. Learn., vol.42, no.3, pp.287-320, 2001.
- [14] Y. Freund, "An adaptive version of the boost by majority algorithm," Mach. Learn., vol.43, no.3, pp.293-318, 2001.
- [15] M. Nakamura, H. Nomiya, and K. Uehara, "Improvement of boosting algorithm by modifying the weighting rule," Ann. Math. Artif. Intell., vol.41, pp.95-109, 2004.
- [16] E.L. Allwein, R.E. Schapire, and Y. Singer, "Reducing multiclass to binary: A unifying approach for margin classifiers," J. Machine Learning Research, vol.1, pp.113-141, 2001.
- [17] 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.
- [18] R.E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-related predictions," Proc. 11th Annual Conference on Computational Learning Theory, pp.80-91, 1998.
- [19] R.C. Nelson, "Finding line segments by stick growing," IEEE Trans. Pattern Anal. Mach. Intell., vol.16, no.5, pp.519-523, 1994.
- [20] B. Leibe and B. Schiele, "Analyzing appearance and contour based methods for object categorization," Proc. International Conference on Computer Vision and Pattern Recognition, pp.409-415, 2003.
- [21] K. Grauman and T. Darrell, "Efficient image matching with distributions of local invariant features," Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.2, pp.627-634, 2005.
- (平成 18 年 11 月 30 日受付, 19 年 5 月 2 日再受付)



野宮 浩揮

平 16 神戸大学大学院自然科学研究科情報知能工学専攻修士課程了。現在、同大学院情報・電子科学専攻博士課程に在学中。機械学習、特に視覚的学習の研究に従事。



上原 邦昭 (正員)

昭 53 阪大・基礎工・情報工学卒。昭 58 同大学院博士後期課程単位取得退学。同産業科学研究所助手、講師、神戸大学工学部情報知能工学科助教授、同都市安全研究センター教授を経て、現在、同大学院自然科学研究科教授。工博。人工知能、特に機械学習、マルチメディア処理の研究に従事。