情報処理学会論文誌 Vol.59 No.8 1499–1510 (Aug. 2018)

クラスタリングとロジスティック回帰を利用した物体概念の学習と認識への応用

秋本 翔平[1,a)] 高橋 智一[2] 鈴木 昌人[2] 青柳 誠司[2]

受付日 2017年10月19日, 採録日 2018年5月10日

概要：精度が高い物体識別方法として CNN（Convolutional Neural Networks）が一般的に利用されている．しかし，この手法は各カテゴリの学習データとして数万枚の画像が必要であり，膨大な学習時間が必要である．またこの手法はどのような特徴を利用して識別を行っているかが分からない．実際に，高い識別率を持つ CNN でも，人間の直感から外れた認識をすることが報告されている．一方で，人間はあるカテゴリの画像を数枚見るだけで，物体の概念を得ることが可能である．さらに人間は言葉で概念を表現することも可能である．本研究では，クラスタリングとロジスティック回帰を利用した物体概念の学習方法を提案する．提案手法は複数の低次元な特徴（具体的には，色，輪郭，大きさ）を利用することで，物体概念の生成を行い，高い識別率と可読性を持つ識別器を作成する．提案手法の有効性を，RGB-D オブジェクトデータセットを用いた従来手法との比較により実証した．

キーワード：物体認識，物体概念，RGB-D センサ，クラスタリング，ロジスティック回帰

Learning of Object Concept and Application to Object Recognition Using Clustering and Logistic Regression

Shohei Akimoto[1,a)] Tomokazu Takahashi[2] Masato Suzuki[2] Seiji Aoyagi[2]

Received: October 19, 2017, Accepted: May 10, 2018

Abstract: In general, CNN (Convolutional Neural Networks) is used as the method with high recognition accuracy. In CNN, however, several tens of thousands images are required as learning data for each category. Also, huge learning time is required. Another drawback is that we cannot understand what CNN focuses the features when recognizing, in fact, it has been reported that well-trained CNN has recognition results that are out of human intuition. In contrast, after a human just look at several objects in a category he can get something like its general object concept. Furthermore, a human can represent the concept by words. In this article, a new concept learning method based on clustering and logistic regression is proposed. The proposed method learns object concepts from multiple low-dimensional features, e.g., color, contour, and size based on which it generates classifier with high accuracy and readability. Effectiveness of the proposed method was demonstrated by comparison with the conventional method using RGB-D Object Dataset.

Keywords: object recognition, object concept, RGB-D sensor, clustering, logistic regression

1. 背景

今後ロボットは工場内のみでなく，家庭内や病院内等で人間と共存しながら様々なタスクを行うことが期待されている．このような環境では，ロボットは動的に変化する環境を自ら認識し，次の行動を計画する必要がある．人間は，部屋の戸口から一瞥するだけでその部屋の環境を瞬時に認識することができる．これは，LRF（Laser Range Finder）のように機械的にフリースペース（物体に占有されておらずロボットが移動可能なスペース）を見つけているのではなく，そこに存在する機器（オフィスならば机，椅子等）の抽出と認識を行い，それから部屋のレイアウトを類推・把

[1] 関西大学大学院理工学研究科システム理工学専攻
Graduate School of Science and Engineering, Kansai University, Suita, Osaka 564–8680, Japan
[2] 関西大学システム理工学部機械工学科
Department of Mechanical Engineering, Kansai University, Suita, Osaka 564–8680, Japan
[a)] k485204@kansai-u.ac.jp

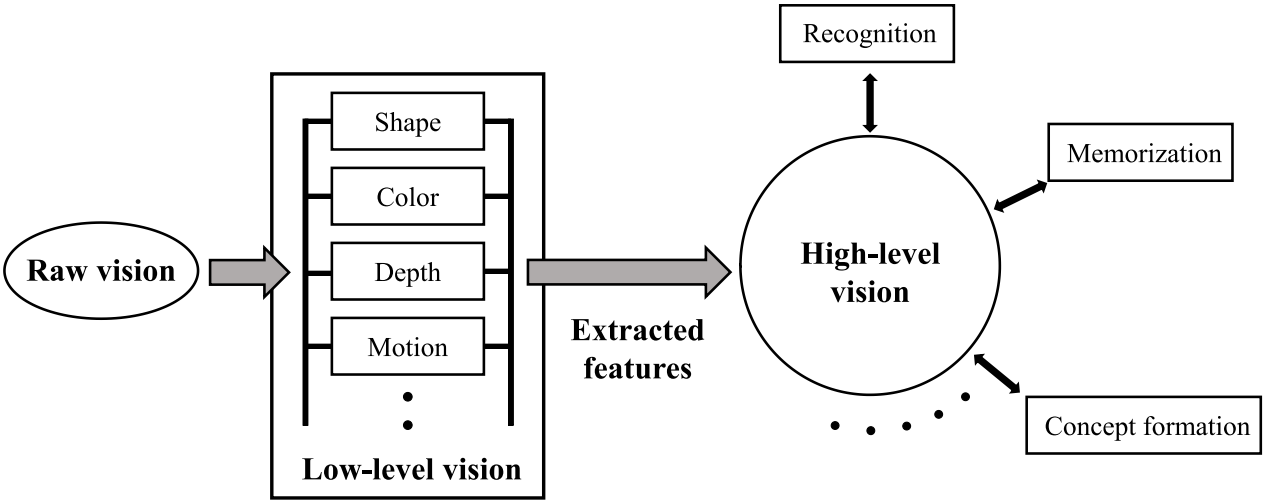

図 1 人間が持つ高次視覚と低次視覚の位置づけ [11]
Fig. 1 High-level vision and low-level vision of human being [11].

握しているものだと考えられる．つまり，環境をリアルタイムに把握するためには，物体認識が重要であるといえる．

従来から行われている物体認識の研究として，大量の画像から得た特徴量を学習することで，学習した物体ごとの識別器を生成するアピアランスベースドな手法と，あらかじめ用意した物体の 3 次元モデルと未知物体の 3 次元データとを比較し，照合を行うことで物体の種類を判別するモデルベースドな手法が存在する．アピアランスベースドな手法では，ウェブカメラや KINECT 等といった画像センサにより獲得した，1 物体につき 1,000 枚といった大量の画像から，Scale-Invariant Feature Transform（SIFT）[1] や Histograms of Oriented Gradients（HOG）[2] といった輝度変化に着目した高次元な特徴量を抽出する．これを用いて，AdaBoost [3] や Support Vector Machine（以下，SVM）[4] 等の学習器により各物体の識別器を生成する．しかし，学習に大量の画像が必要なため，学習に時間がかかること，特徴量が高次元であるため，誤認識した理由が分かりにくいという問題がある [5]．モデルベースドな手法 [6] では，3 次元データを取り扱うため，計算コストが高く認識に時間がかかるという問題がある．

最近では，Deep Learning と呼ばれる大規模データを多層化したニューラルネットにより学習することで，特徴点の抽出から識別器の生成までを自動的に行う手法が提案されている [7]．しかし，アピアランスベースドな手法よりも，さらに高次元で大量の学習データが必要であり，学習に要するコストも高い．学習枚数が大量に必要となる問題に対して，Sachin らはメタラーナ（CNN を監視する上位のネットワーク）を用いて CNN のパラメータの学習を行うモデルを提案している [8]．Sachin らのモデルでは，少ない学習枚数で高い識別率を達成しているが，識別できるカテゴリの数が 5 種類と少なく，数十種類のカテゴリに対応できるかどうかはまだ判明していない．また，Deep Learning では，学習の過程において物体のどのような特徴に着目しているかが分からないという問題点がある．実際に，Nguyen らによって，高い識別率を持つ畳み込みニューラルネットワークが，人間の直感から外れた認識を行うことがあると報告されている（たとえば，ペンギンの画像は正しく認識されるが，プログラムによって生成された無意味な幾何学模様の画像も，ペンギンと認識してしまう）[9]．これについては，5.2 節で議論をする．

一方で，人間の認知行動について Ullman や横澤らは，**図 1** に示すように，物体認識や記憶，概念生成は高次視覚と呼ばれる複雑な視覚情報処理の一種であると定義しており，これらの高次視覚は，形状や色，奥行き等の低次視覚と呼ばれる認識処理の組合せによって構成されていることを報告している [10], [11]．Muller らは，認識を行う際に，認識対象についての事前知識によって，特定の低次視覚が重み付けされることを報告している [12]．また，人間が行う物体認識では，従来の機械学習と比較して，少ない枚数の画像からその物体が何であるかを判断することが可能であり，なぜその物体と判断したのかの理由を説明することが可能である．これらの研究から，人間は既知の物体に対して，形状や色のような低次視覚から形成される物体に固有な概念（物体概念）を持っており，既知の物体を見たときに，物体概念に基づいた認識を行っているものと考える．

以上の推察に基づき，本稿では，人間の行っている認知行動を参考に，色，輪郭線，大きさといった特徴（低次視覚）を適切な重みにより統合したものを物体概念（高次視覚）と定義したうえで，数少ない学習データを利用した物体概念の学習方法とそれに基づいた一般物体の識別方法を提案する．51 種類のカテゴリを含む RGB-D Object Dataset [13] から取り出した学習データにより物体概念の学習を行う．これに基づいた識別結果により物体概念の妥当性を判断する．具体的な判断基準として，(1) 物体概念

に基づいた識別器が，高い精度を持っていること，(2) 物体概念に基づいた識別結果から得られる上位のカテゴリを比較した際に，色，輪郭線，大きさから構成される視覚的な共通点が確認できること，(3) 学習した物体概念，すなわち，各特徴の重みから，その物体の色，形状，大きさが推察できること，以上の 3 点を設けることで，学習された概念が適切か否かを判断する．

本稿の構成を述べる．2 章で，物体概念を記述する画像特徴の計算方法について述べる．3 章で，クラスタリングとロジスティック回帰を利用した物体概念の学習と識別方法について，その処理手順と生成される概念について述べる．4 章で，提案手法の評価に使用するデータセットの詳細とその評価方法について述べる．5 章で，提案手法の評価実験の結果について述べる．6 章で，結論と今後の課題について述べる．

2. 物体概念を記述する特徴

提案手法では，物体概念を表現する特徴として，物体の色，輪郭，大きさを利用する．色を表現する特徴として，ファジィカテゴリカルカラーヒストグラム（以下，FCCH）[14] を利用する．物体の輪郭を表現する特徴として，P 型フーリエ記述子 [15] を利用する．物体の大きさを表現する特徴として，縦方向の長さと横方向の長さを利用する．

2.1 ファジィカテゴリカルカラーヒストグラム

FCCH は，Han らが提案したファジィカラーヒストグラム（FCH）[14] の階級（ヒストグラムのビン）に，内川の提案した 13 色のカテゴリカルカラーを適用した特徴である [16]．通常のカラーヒストグラムは，入力された色と最も近い色の階級に投票する．FCH ではカテゴリカルカラーをクラスタと見なしたファジィクラスタリング [17] により各ピクセルの帰属度を計算し，得られた帰属度をそのまま投票する．階級の境界が曖昧になるため，元画像の色情報を損なうことがない．カテゴリカルカラーとは，日本人を被験者とした実験的検証によって得られた 13 色のことであり，人間はすべての色をこの 13 色に分類しやすいことが知られている．これをヒストグラムの階級にすることで，人間（日本人）に近い色の表現ができるようになる．

FCCH の計算過程を以下に示す．入力画像の各画素は RGB 値で与えられているものとする．

I. 入力画像の各画素値を RGB 色空間から CIEL*a*b*色空間へと変換する [18]．
II. 各画素において，カテゴリカルカラーをクラスタとして，式 (1) から入力データの i 番目のクラスタへの帰属度 u_i を計算する（$i = 1 \sim 13$）．

$$u_i = \frac{1}{\left[\sum_{k=1}^{13} \frac{d(\boldsymbol{x}-\boldsymbol{c}_i)}{d(\boldsymbol{x}-\boldsymbol{c}_k)}\right]^{\frac{1}{m-1}}} \tag{1}$$

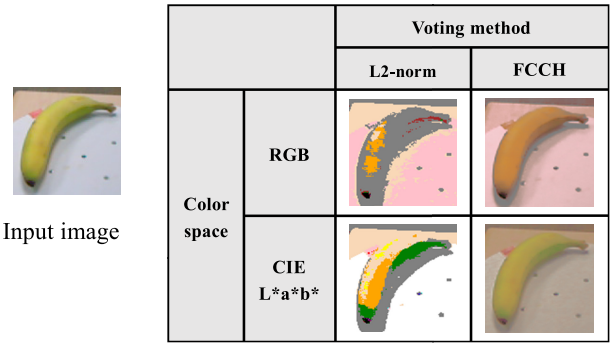

図 2 各画素の 13 色カテゴリカルカラーへの色置換結果
Fig. 2 Results of transforming color value of each pixel to 13 categorical colors in various methods.

ここで，$\boldsymbol{x}$ は入力データ，$\boldsymbol{c}_i$ はクラスタ中心，m は曖昧さを表すファジィ度であり，値が大きくなるほど，曖昧さが大きくなる．m の範囲は $1 < m$ である．$d(\boldsymbol{x}-\boldsymbol{c}_i)$ は入力データとクラスタ中心の差を表している．本節では，$\boldsymbol{x}$ が各画素から得られる CIEL*a*b*値を，$\boldsymbol{c}_i$ が 1 つのカテゴリカルカラーに対応している．また $d(\boldsymbol{x}-\boldsymbol{c}_i)$ については，CIEDE2000 色差式 [19] によって計算する．

III. 得られた各階級の帰属度の総和を計算し，画素数で割り，正規化する．

図 2 に RGB 色空間と CIEL*a*b*色空間の画像に対し，各ピクセルを L2 ノルムが最も近いカテゴリカルカラーに置換した結果と，FCCH により得られた帰属度を用いて 13 色のカテゴリカルカラーの重み付けをすることで合成された色に置換した結果を示す．RGB 値と L2 ノルムによるカテゴリカル色の置換結果（図 2 左上）では，多くの色が灰色に置換されているが，CIEL*a*b*値と FCCH による置換結果（図 2 右下）では，これが改善されていることが分かる．すなわち，FCCH では，従来のカラーヒストグラムと比較して，画像の情報を良好に特徴量化できていることが分かる．特徴量の次元は，13 次元となる．

2.2 P 型フーリエ記述子

P 型フーリエ記述子は，物体の輪郭線の全曲率関数を指数部に持つ複素関数をフーリエ展開して得られる記述子である [15]．記述子はスケール変化や平行移動に対し不変であり，その低域成分には輪郭線の形状に関する情報が多く含まれている．

以下に，P 型フーリエ記述子の算出法を述べる．**図 3** に示すように，画像平面上の輪郭線を複素平面上の点の集まりと見なす．次に，輪郭線上の始点から終点まで，1 辺の長さが δ になるように，新しく点を置き直すことで，n 本の長さ δ の線分からなる折線図形に近似する．各点を $z(j)$ として，各線分を δ で正規化した $w(j)$ は式 (2) のようになる．

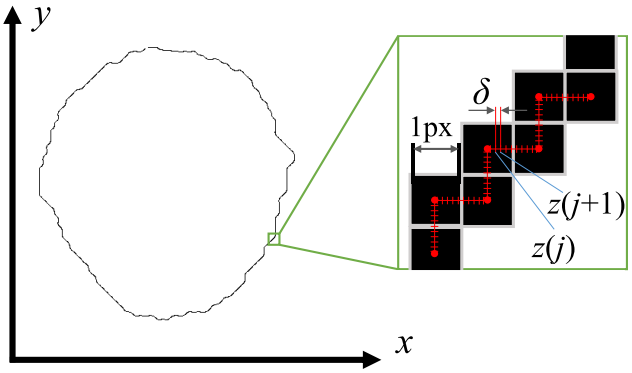

図 3 輪郭線からの P 型フーリエ記述子の計算方法
Fig. 3 Obtaining method of P-type Fourier descriptor from contour.

$$w(j) = \frac{z(j+1) - z(j)}{\delta} \tag{2}$$

$w(j)$ を離散フーリエ変換することで，係数 $c(k)$ を得る．

$$c(k) = \frac{1}{n}\sum_{j=0}^{n-1} w(j) \exp\left(-2\pi i \frac{jk}{n}\right) \tag{3}$$

ここで，$k = -(N-1), \ldots, 0, \ldots, N$ である．式 (3) で算出される係数の集合が N 次の P 型フーリエ記述子となる．経験的に $N = 16$ とし，フーリエ展開したスペクトルを特徴量とする．特徴量の次元は，$2 \times N = 2 \times 16 = 32$ 次元となる．

2.3 大きさ

大きさは，物体の距離画像から縦方向の長さと横方向の長さを特徴として抽出する．単位はメートルとする．物体が回転することを考慮して，縦方向と横方向の長さの中で，大きい方を長手方向の長さ，小さい方を短手方向の長さと定義する．これらの長さの特徴を 2 次元にまとめた場合，短手方向の長さはつねに長手方向の長さより小さくなるため，ユークリッド距離等の計算を行った場合，短手方向の長さが無視されやすくなることを予備的な実験で確認した．そのため，本研究では長手方向の長さと短手方向の長さをそれぞれ 1 次元の特徴として扱う．大きさは正規化せずに扱う．

3. 物体概念の学習と識別について

本章では，FCCH，P 型フーリエ記述子，長手方向の大きさ，短手方向の大きさの 4 つの特徴を利用した物体の概念の学習と識別の方法について述べる．本研究では，4 つの特徴ごとに，後述する RGB-D Object Dataset [13] に含まれる 51 種類のカテゴリの中で，どのカテゴリに類似しているかを表す「カテゴリ帰属度」と，4 つの特徴と 51 種類のカテゴリの $4 \times 51 = 204$ 個の帰属度の中で，どの帰属度を重要視するかを表す「カテゴリ帰属度の重み」を用いて，物体の概念を表現する．以下に，「カテゴリ帰属度」と「カテゴリ帰属度の重み」について述べる．

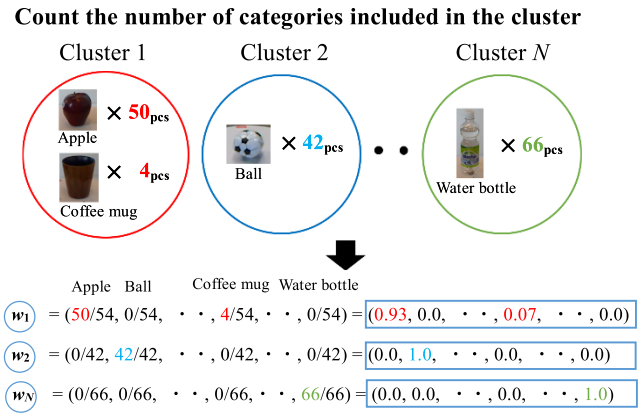

図 4 各クラスタに含まれるカテゴリの割合の計算方法
Fig. 4 Procedure for calculating the ratio of categories included in each cluster.

3.1 カテゴリ帰属度について

カテゴリ帰属度は，識別対象となる物体が，51 種類のカテゴリの中で，どのカテゴリに類似しているかを表すヒストグラムである．カテゴリ帰属度は 4 つの特徴のそれぞれで独立して計算する．1 つの識別対象から得られるカテゴリ帰属度は，$4 \times 51 = 204$ 個の帰属度で構成される．本節では，カテゴリ帰属度を計算する際に必要となる学習と，カテゴリ帰属度の計算方法について述べる．

3.1.1 クラスタリングによる学習について

学習データから得られる FCCH，P 型フーリエ記述子，長手方向の大きさ，短手方向の大きさに対し，それぞれ Affinity Propagation [20]（以下，AP）によるクラスタリングを行う．AP は入力された全データ間の類似度をもとに，推定したクラスタ中心とそのクラスタに属するデータ間の類似度を最大にするクラスタリング手法である．また，クラスタ数の推定まで行うことができるため，ユーザがクラスタ数を設定する必要がない．学習データに対し AP によるクラスタリングを行うことで，類似した特徴をまとめることができ，概念を生成しやすくなるものと考える．

AP によるクラスタリングによって，N 個のクラスタが生成された場合，**図 4** のように各クラスタにはそれぞれ異なる個数で学習データのカテゴリが含まれる．この個数を用いて，各クラスタに含まれる 51 種類のカテゴリの割合を求める．このとき，i 番目のクラスタが持つ 51 種類のカテゴリの割合を $\boldsymbol{w}_i$（ここで，$i = 1, \ldots, N$ である）と定義する．$\boldsymbol{w}_i$ 自体の次元は 51 次元である．本研究では，N 個のクラスタ中心 $\boldsymbol{c}_i$ とカテゴリ割合 $\boldsymbol{w}_i$ を求めることをクラスタリングによる教師あり学習と見なす．

3.1.2 カテゴリ帰属度の計算方法

識別対象であるデータが，3.1.1 項で得られるクラスタの中でどれに類似しているかを，2.1 節と同様にファジィクラスタリングの帰属度 u_i によって表現する．

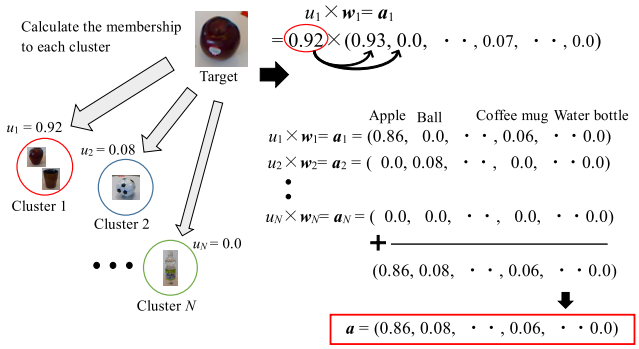

図 5 カテゴリ帰属度の計算手順

Fig. 5 Procedure for calculating the categories of membership.

$$u_i = \frac{1}{\left[\sum_{k=1}^{N} \frac{(\boldsymbol{x}-\boldsymbol{c}_i)^2}{(\boldsymbol{x}-\boldsymbol{c}_k)^2}\right]^{\frac{1}{m-1}}} \tag{4}$$

ここで，$\boldsymbol{x}$ は識別対象から抽出した特徴ベクトル，$\boldsymbol{c}_i$ は AP によって得られたクラスタ中心，m はファジィ度である．式 (4) によって得られる帰属度 u_i はクラスタへの帰属度を表しているため，このままではどのカテゴリに類似しているか判断できない．そこで，**図 5** のように，3.1.1 項で得られた各クラスタのカテゴリ割合 $\boldsymbol{w}_i$ を用いてカテゴリ帰属度 $\boldsymbol{a}$ に変換する．以下に，カテゴリ帰属度 $\boldsymbol{a}$ への変換方法を示す．

1. 各クラスタの帰属度 u_i とカテゴリ割合 $\boldsymbol{w}_i$ を掛けることで $\boldsymbol{a}_i$ を得る．i は 1 から N まで計算する．
2. 得られた計算結果を，$\boldsymbol{a}_1$ から $\boldsymbol{a}_N$ まで同じ成分（カテゴリに相当）で足し合わせ，カテゴリ帰属度 $\boldsymbol{a}$ を得る．

カテゴリ帰属度 $\boldsymbol{a}$ の次元は 51（カテゴリ数）である．$\boldsymbol{a}$ を FCCH，P 型フーリエ記述子，長手方向の大きさ，短手方向の大きさの 4 つの特徴ごとに得る．

3.2 カテゴリ帰属度の重み

カテゴリ帰属度の重みは，3.1 節で述べた 4 つの特徴と 51 種類のカテゴリの $4 \times 51 = 204$ 個の帰属度の中で，どれが重要であるかを表すヒストグラムである．重みはカテゴリごとに求められる．たとえば，Apple のカテゴリを識別する際には，Apple の色や形状に類似したカテゴリに大きな重みを与え，Water bottle のカテゴリを識別する際には，Apple とは別の，Water bottle の色や形状に類似したカテゴリに大きな重みを与える．本節では，カテゴリ帰属度の重みの学習方法と，カテゴリ帰属度とその重みを合わせた識別方法について述べる．

3.2.1 カテゴリ帰属度の重みの学習について

3.1.1 項で AP のクラスタリングに利用した学習データから，カテゴリ帰属度を 4 つの特徴ごとに計算する．このとき，1 つの学習データから得られるカテゴリ帰属度は，FCCH の特徴から得られる $\boldsymbol{a}^{FCCH}$，P 型フーリエ記述子から得られる $\boldsymbol{a}^{Fourier}$，長手方向の大きさから得られる $\boldsymbol{a}^{Size(long)}$，短手方向の大きさから得られる $\boldsymbol{a}^{Size(short)}$ の計 4 つある．このカテゴリ帰属度を水平に結合して，式 (5) のようにカテゴリ帰属度 $\boldsymbol{h}$ を得る．$\boldsymbol{h}$ の次元は，$4 \times 51 = 204$ 次元になる．

$$\boldsymbol{h} = [\boldsymbol{a}^{FCCH}, \boldsymbol{a}^{Fourier}, \boldsymbol{a}^{Size(long)}, \boldsymbol{a}^{Size(short)}] \tag{5}$$

すべての学習データにおいて，各特徴から得られるカテゴリ帰属度を結合して，カテゴリ帰属度の重みの学習データとする．

学習データとロジスティック回帰 [21] を用いて，カテゴリ帰属度 $\boldsymbol{h}$ における各成分（各カテゴリ）における重みを学習する．ロジスティック回帰により，式 (6) を最大にするようなパラメータ $\boldsymbol{\theta}$，$\boldsymbol{b}$ を求める．

$$\max_{\boldsymbol{\theta},\boldsymbol{b}} \left[\frac{1}{n} \sum_{i=1}^{n} \log \frac{e^{(\boldsymbol{\theta}^{(y(i))}\boldsymbol{h}^{(i)} + \boldsymbol{b}^{(y(i))})}}{\sum_{c \in y} e^{(\boldsymbol{\theta}^{(c)}\boldsymbol{h}^{(i)} + \boldsymbol{b}^{(c)})}} - \frac{1}{2} \sum_{c \in y} \|\boldsymbol{\theta}^{(c)}\|^2 \right] \tag{6}$$

ここで，n は学習データ数，$\boldsymbol{\theta}^{(c)}$（204 次元）は c 番目のカテゴリの帰属度の重み，$\boldsymbol{\theta}^{(y(i))}$ は学習データが所属するカテゴリ（正解カテゴリ）の帰属度の重み，$\boldsymbol{b}^{(c)}$ は c 番目のカテゴリの重みの定数項，$\boldsymbol{b}^{(y(i))}$ は i 番目の学習データが所属するカテゴリ（正解カテゴリ）の重みの定数項，$\boldsymbol{h}^{(i)}$ は i 番目の学習データのカテゴリ帰属度である．

式 (6) を解くことで，各カテゴリの帰属度の重みである $\boldsymbol{\theta}$ を求めることができる．本研究では，51 種類のカテゴリごとに，204 次元のカテゴリ帰属度 $\boldsymbol{h}$ の重みを求めることになるため，学習によって得られる $\boldsymbol{\theta}$ のパラメータは，$51 \times 204 = 10{,}404$ 個になる．

3.2.2 カテゴリ帰属度と重みを用いた識別方法について

識別対象であるデータから，カテゴリ帰属度 $\boldsymbol{a}$ を 4 つの特徴ごとに求め，水平に結合することで，カテゴリ帰属度 $\boldsymbol{h}$ を得る．この $\boldsymbol{h}$ と 3.2.1 項で得られたカテゴリ帰属度の重み $\boldsymbol{\theta}$，カテゴリの重みの定数項 $\boldsymbol{b}$ を用いて，識別対象が y 番目のカテゴリであることを表す確率 $p(y \mid \boldsymbol{h})$ を式 (7) のように計算する．

$$p(y \mid \boldsymbol{h}) = \frac{e^{(\boldsymbol{\theta}^{(y)}\boldsymbol{h} + \boldsymbol{b}^{(y)})}}{\sum_{c \in y} e^{(\boldsymbol{\theta}^{(c)}\boldsymbol{h} + \boldsymbol{b}^{(c)})}} \tag{7}$$

式 (7) は，識別対象から得られるカテゴリ帰属度 $\boldsymbol{h}$ と各カテゴリの重み $\boldsymbol{\theta}$ を掛け合わせた内積の結果を soft-max 関数で正規化したものである．得られた $p(y \mid \boldsymbol{h})$ の中で，最も値が大きかった y を識別対象が属するカテゴリと見なす．

3.3 提案手法の特徴について

提案手法の特徴を**図 6** に示す．図 6 は，提案手法の特徴を説明するために，例として，Apple，Comb，Marker

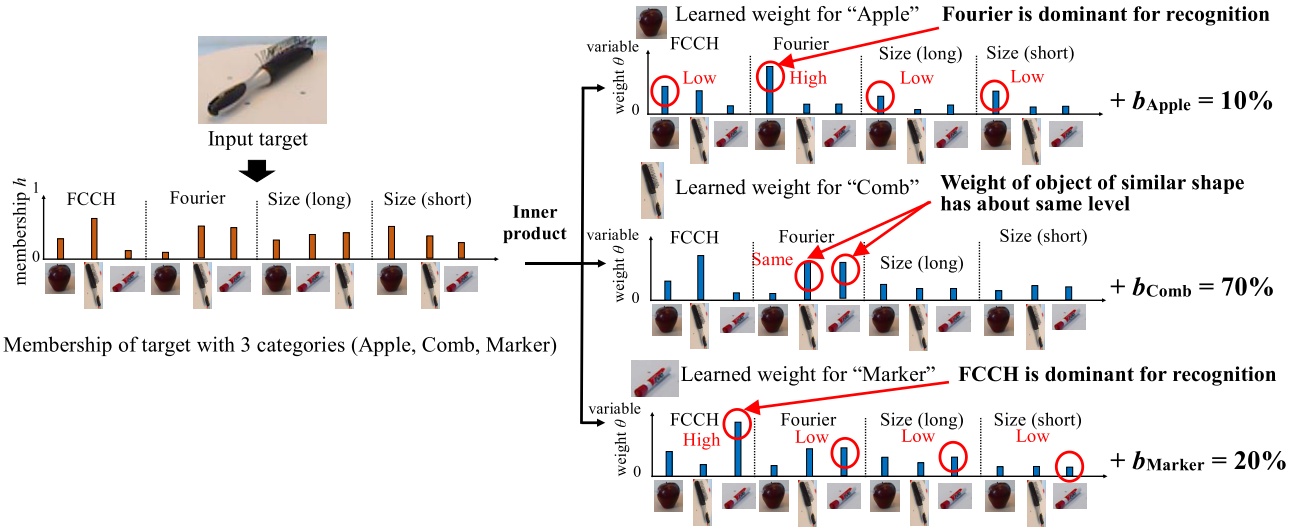

図 6 提案手法における認識の流れの概念図．例として，3 つのカテゴリ（Apple，Comb，Marker）と 4 つの特徴（FCCH，Fourier，Size (long)，Size (short)）で説明している

Fig. 6 Schematic of flow of object recognition by proposal method. An example of 3 categories (Apple, Comb, Marker) and 4 features (FCCH, Fourier, Size (long), Size (short)).

の 3 つのカテゴリと，FCCH，P 型フーリエ記述子，長手方向の大きさ，短手方向の大きさの 4 つの特徴から得られるカテゴリ帰属度とカテゴリ重みを仮定して，提案手法の特徴を説明している概念図である．

入力された識別対象は Comb が正解のカテゴリである．識別対象から得られるカテゴリ帰属度は，FCCH では正しく Comb のカテゴリに大きな帰属度が表れているが，Fourier や Size(long) では Comb と Marker のどちらか判断しにくい帰属度の値になっており，Size(short) では間違ったカテゴリに高い帰属度が表れている．一方で，Comb の帰属度の重みを見ると，P 型フーリエ記述子から得られるカテゴリ帰属度の中でも，Comb と Marker に大きな重みを持っていることが分かる．

実際に，式 (7) を用いて識別した場合では，識別対象から得られるカテゴリ帰属度と Comb の重みから得られる内積の計算結果が最大となるため，識別対象は Comb と判定される．この過程について以下に説明する．

前述したように，識別対象から得られる P 型フーリエ記述子のカテゴリ帰属度の中で，Comb と Marker は類似した輪郭線を持つため，同程度の帰属度を持っている（図 6 左）．次に重みの方を考える（図 6 右）．Comb は，P 型フーリエ記述子において細長い棒状の物体に対して大きな重みを持っており，細長い物体であることが Comb の概念として重要であることが分かる．一方，今回は赤い Marker を対象としたので，FCCH において赤い物体に対して大きな重みを持っており，これが Marker の識別に重要であることが分かる．以上のような状況なので，Comb の画像が入力された場合，カテゴリ帰属度（図 6 左）と重み（図 6 右）とで内積を計算すると，Comb が最尤の認識対象として出力されることになる．このように，学習によって得られる重み $\boldsymbol{\theta}$ が物体を特徴づける概念を表現していることになる．

図 7 RGB-D Object Dataset [13]

Fig. 7 RGB-D Object Dataset [13].

4. 提案手法の評価の手順

本章では，提案手法の評価に利用するデータセットである RGB-D Object Dataset [13] と提案手法の評価方法について述べる．

4.1 RGB-D Object Dataset

RGB-D Object Dataset は，**図 7** に示すような，家庭でよく見られる 51 カテゴリ 300 種類の物体を扱っている．すなわち，同じカテゴリに複数の種類がある．たとえばリンゴのカテゴリに赤いリンゴ，青いリンゴ等の種類がある．RGB-D Object Dataset は，これらの物体に対するカラー

表 1 各カテゴリのデータ数
Table 1 Number of data of each category.

category name	number of data	category name	number of data	category name	number of data
Apple	607	Food can	822	Orange	710
Ball	783	Food cup	786	Peach	674
Banana	699	Food jar	768	Pear	674
Bell pepper	634	Garlic	780	Pitcher	549
Binder	710	Glue stick	801	Plate	712
Bowl	806	Greens	728	Pliers	590
Calculator	632	Hard towel	783	Potato	634
Camera	573	Instant noodles	771	Rubber eraser	784
Cap	668	Keyboard	709	Scissors	625
Cell phone	544	Kleenex	804	Shampoo	783
Cereal box	570	Lemon	593	Soda can	628
Coffee mug	555	Lightbulb	578	Sponge	631
Comb	558	Lime	630	Stapler	595
Dry battery	567	Marker	806	Tomato	681
Flashlight	605	Mushroom	778	Toothbrush	595
Food bag	788	Notebook	798	toothpaste	813
Food box	787	Onion	792	Water bottle	561

画像，距離画像，物体領域のマスク画像を含むデータセットである．物体の撮影には KINECT を用いており，カラー画像と距離画像が同時に取得されている．撮影の際は，物体をターンテーブルに設置し，物体から 1.0 m 離れた場所から 3 段階の高さで物体 1 周分のデータを取得している．

本稿では，RGB-D Object Dataset の 51 カテゴリの各カテゴリから，図 7 の画像で示す 1 種類を取り出し，これに対する多数の画像を学習用と識別用に分けて利用する．各カテゴリのデータ数を**表 1** に示す．実験に用いる学習用のデータセットと識別用のデータセットの詳細を以下に記述する．学習枚数を変化させたときの認識結果を検討するために，学習枚数は各カテゴリ 10 枚ずつから 100 枚ずつまで，10 枚刻みで変化させる．学習と識別に用いる画像はつねに被らないようにする．たとえば，総数が 607 枚である Apple のカテゴリを 100 枚の画像で学習させた場合，識別対象となる画像は残りの $607 - 100 = 507$ 枚となる．

4.2 学習方法について

本研究では，3.1.1 項におけるカテゴリ帰属度を計算するための教師あり学習と，3.2.1 項におけるカテゴリ帰属度の重みを得るためのロジスティック回帰による学習の，2 通りの学習がある．

カテゴリ帰属度を計算するための学習では，各枚数の学習データから，FCCH，P 型フーリエ記述子，長手方向の大きさ，短手方向の大きさの 4 つの特徴を抽出し．AP によるクラスタリングを行い，それぞれの特徴で，クラスタ中心 $\boldsymbol{c}_i$ とカテゴリ割合 $\boldsymbol{w}_i$ を学習する．

カテゴリ帰属度の重みを得るための学習では，クラスタリングに用いたデータをそのまま利用して，4 つの特徴それぞれでカテゴリ帰属度を求め，水平に結合し，カテゴリ帰属度 $\boldsymbol{h}$ を求める．得られた $\boldsymbol{h}$ を式 (6) に代入し，最適化することで，カテゴリ帰属度の重み $\boldsymbol{\theta}$ を学習する．カテゴリ帰属度を求める際に，式 (4) 中のファジィ度 m を 1.1 から 2.0 まで 0.1 刻みで変化させ，ファジィ度の変化による認識結果を検討する．

4.3 識別について

識別では，学習と同様に，識別用データから 4 つの特徴を抽出し，学習によって得られたクラスタ中心 $\boldsymbol{c}_i$ とカテゴリ割合 $\boldsymbol{w}_i$ から，カテゴリ帰属度を計算する．このとき，ファジィ度 m は学習時と同じ値に設定する．4 つの特徴から得られたカテゴリ帰属度を水平に結合してカテゴリ帰属度 $\boldsymbol{h}$ とし，式 (7) を用いて，51 種類のカテゴリに属する確率 $p(y\,|\,\boldsymbol{h})$ を計算する．

識別結果の評価には，$p(y\,|\,\boldsymbol{h})$ の中で，最も高い値を持つカテゴリと識別用のデータのラベルが一致する割合（Top-rate）と，識別用のデータのラベルが $p(y\,|\,\boldsymbol{h})$ の上位 5 位以内に含まれる割合（Ranking-rate）とを用いる．

4.4 評価方法

比較対象として，① Histogram of Oriented Normal Vector [22]（以下，HONV）を特徴として，SVM [4] によって作

成された識別器（以下，HONV＋SVM），② CNN（Convolutional Neural Networks）のモデルの一種である AlexNet [7] による識別器，③ DeCAF [23] 特徴を入力とする SVM [4] によって作成された識別器（以下，DeCAF＋SVM）の 3 種類と比較する．AlexNet と DeCAF＋SVM は，CNN のライブラリである Caffe [24] を用いて作成する．

HONV＋SVM では，RGB-D Object Dataset に含まれる距離画像から，HONV を抽出し，SVM によって学習することで識別器を作成する．学習枚数は，提案手法と同様に，学習枚数を 10 枚から 100 枚まで 10 枚刻みで変化させて比較を行う．学習の際には，SVM のコストパラメータを変化させ，最も識別率が高かったものを識別結果として採用する．SVM では，識別結果が確率で取得できないため，Top-rate のみで比較を行う．

CNN では，AlexNet を用いて，RGB-D Object Dataset に含まれるカラー画像のみを用いて学習させる．学習枚数は，各カテゴリ 10 枚と 100 枚，つまり，$51 \times 10 = 510$ と $51 \times 100 = 5{,}100$ の学習枚数で比較を行う．AlexNet における学習では，モデルの重みをランダムに初期化する方法と，ImageNet データセット [25] ですでに学習済みである重みからファインチューニングを行う方法の 2 通りの方法で比較を行う [26]．CNN の学習パラメータは文献 [26] を参考に設定を行い，CNN の最適化手法には Stochastic Gradient Descent（SGD）を利用する．学習率は初期値を 0.001 に設定し，エポック数が 392 回を経るごとに，学習率を 0.1 倍ずつ減少させる．最大エポック数は 1,960 回とする [26]．

DeCAF＋SVM では，RGB-D Object Dataset に含まれるカラー画像から DeCAF を抽出し，SVM によって学習することで識別器を作成する．DeCAF の抽出は，ImageNet データセット [25] ですでに学習済みである AlexNet を利用し，入力画像をネットワークに与えた後，1 番目の全結合層から得られる 4,096 次元の出力を取り出すことで行う [23]．学習枚数は，10 枚から 100 枚まで 10 枚刻みで変化させて比較を行う．学習の際には，SVM のコストパラメータを変化させ，最も識別率が高かったものを識別結果として採用する．SVM では識別結果が確率で取得できないため，Top-rate のみで比較を行う．

5. 実験結果

5.1 識別率による提案手法と他手法の比較

表 2 に各手法において最も識別率が高かった結果を示す．**図 8** に各手法において，学習枚数を変化させた場合の識別率の推移を示す．ここではパラメータ（ファジィ度 m，コストパラメータ cost）を変化させて最も高い識別率が得られた際のデータを示す．

提案手法では，学習枚数が各カテゴリで 100 枚かつファジィ度 m が 1.4 のときに最も高い識別率が得られた．

表 2 評価結果
Table 2 Evaluation results.

	Top-rate	Rank-rate
Proposal method	98.8%	99.9%
HONV+SVM	87.2%	
AlexNet	99.8%	99.9%
AlexNet (Fine-tuning)	99.9%	99.9%
DeCAF+SVM	99.9%	

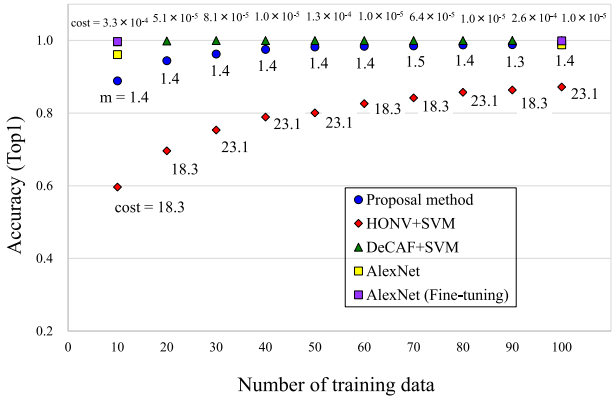

図 8 学習枚数を変化させたときの Top-rate の推移
Fig. 8 Top-rate by changing number of training data.

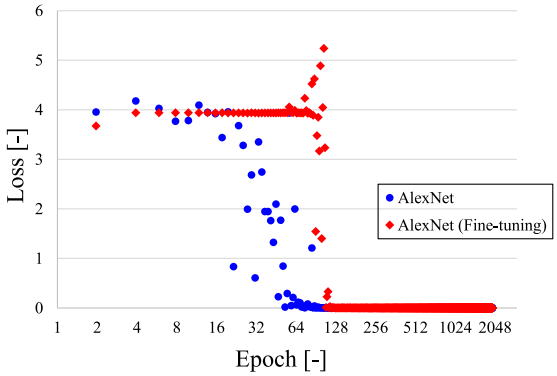

図 9 AlexNet の学習における loss-epoch グラフ
Fig. 9 The graph of loss-epoch in training of AlexNet.

HONV＋SVM では，学習枚数が各カテゴリで 100 枚かつコストパラメータが 23.1 のときに最も高い識別率が得られた．AlexNet によって作成された識別器では，モデルの重みをランダムに初期化する方法とファインチューニングを行う方法の両方で，各カテゴリで 10 枚と 100 枚の画像を用いて学習させた結果を記載している．また，100 枚の学習データを AlexNet に与えた際に，学習中に得られるエポック数に対する損失のグラフを**図 9** に示す．AlexNet による 2 つの手法で，損失が 1.0×10^{-4} 以下の値になり，学習が適切に収束していることが分かる [26]．なお，学習枚数が 10 枚の場合でも，同様に学習が収束することを確認している．2 つの手法ともに，学習枚数が 10 枚の場合でそれぞれ 96%，99%，学習枚数が 100 枚の場合でともに 99%と，高い識別率が得られた．DeCAF＋SVM では，学習枚数が 90 枚かつコストパラメータが 2.6×10^{-4} のとき

に最も高い識別率が得られた．

識別結果を比較すると，提案手法の識別率は，HONV＋SVM のような高次元のヒストグラム特徴を利用した識別手法と比較して 98.8%と高く，提案手法により学習されたカテゴリ重み（物体概念）による識別が有効であること，つまり，識別に適した物体概念が生成されていることを示している．

一方で，提案手法と CNN を利用した手法である AlexNet や DeCAF＋SVM による識別率を比較すると，後者は少ない学習枚数で，提案手法より高い識別率を達成している．この理由は，CNN を用いて識別器を作成する場合に一般的に多くの学習枚数が必要となるが，本実験のデータセットでは物体のカテゴリが 51 種類と少ないため，比較的容易に学習が行えたためだと考える．これらの結果から，識別率の比較だけでは，CNN を用いた手法に対する提案手法の優位性は確認できない．

ただし，提案手法でも学習枚数が 10 枚で 89%の高い識別率が得られており，さらに学習枚数 100 枚で 98.8%と CNN を用いた手法に対して識別率の差を 1.1%に縮めることを達成している．このことは逆説的に，今回採用した色，形状，大きさの特徴が，物体の概念形成に高く寄与していることを示している．CNN を用いた手法は後述するように識別器において物体の概念が形成されていたとしてもその可読性が低い（すなわち，どの特徴が認識に寄与しているのかがよく分からない）．CNN では，おそらくは本提案手法で採用した色，形状，大きさだけでなくその他の多くの特徴をすべて取り入れた総体としての概念が形成されているものと思われ，このため今回の実験では 100%に近い高い識別率が得られたものと考える．

5.2 識別結果のカテゴリによる提案手法と他手法の比較

図 10 に提案手法と CNN である AlexNet のうち重みをランダムに初期化した方法，同じくファインチューニングによる手法の 3 種類の手法による，Ranking-rate の上位 3 つの識別結果の例を示す．図 10 に記載されている数値は，入力されたデータが識別結果のカテゴリに所属する確率である．

提案手法の識別結果を考察する．(1) Camera の入力画像は，黒味のある色と四角い形状をしている．提案手法の Camera の識別結果では，Camera，Stapler，Calculator を上位 3 つの識別結果にあげており，すべての識別結果に黒味のある色と四角い形状という視覚的な共通点があることが確認できる．(2) Food cup の入力画像は銀，白，青色が含まれている円柱である．提案手法の Food cup の識別結果では，Food cup，Food can，Rubber eraser を上位 3 つの識別結果としている．ここで，識別確率が 1%を下回る Rubber eraser を無視した場合，Food cup と Food can には，銀色の蓋，青と白色のラベル，円柱状といった視覚的な共通点を持つことを確認できる．(3) Marker の入力画像は，赤と白色を含む細長い棒状の物体である．提案手法の Marker の識別結果では，Marker，Toothbrush，Mushroom を上位 3 つの識別結果としており，識別確率が 1%を下回る Mushroom を無視した場合，Marker と Toothbrush には，赤と白色からなる細長い棒状の物体という視覚的な共通点が確認できる．

提案手法から得られた結果を見ると，上位の識別結果には，入力したカテゴリが持つ色や形状において，多くの共通点があることが分かる．これは，提案手法が物体を識別する際に，各カテゴリに固有な特徴を学習できており，本稿で定義した色，輪郭線，大きさにより構成される物体概念が生成されていることが分かる．

次に，CNN を用いた手法である AlexNet の識別結果を考察する．CNN の構造の一部は，哺乳類の初期視覚野の

Input data	Results of Proposed method			Results of AlexNet			Results of AlexNet (Fine-tuning)		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
 Camera	 Camera 0.40	 Stapler 0.33	 Calculator 0.27	 Camera 0.99	 lime 5.3×10^{-5}	 Stapler 9.8×10^{-9}	 Camera 0.99	 lime 9.4×10^{-5}	 Stapler 3.1×10^{-5}
 Food cup	 Food cup 0.55	 Food can 0.44	 Rubber eraser 2.0×10^{-3}	 Food cup 0.99	 Pliers 7.6×10^{-8}	 Kleenex 5.2×10^{-9}	 Food cup 0.99	 Food can 1.3×10^{-5}	 Food jar 5.2×10^{-6}
 Marker	 Marker 0.59	 Toothbrush 0.22	 Mushroom 0.08	 Marker 0.99	 Toothpaste 4.4×10^{-6}	 Scissor 2.8×10^{-7}	 Marker 0.99	 Greens 1.2×10^{-3}	 Toothbrush 8.5×10^{-4}

図 10 提案手法と AlexNet から得られる識別率

Fig. 10 Recognition probability of each example.

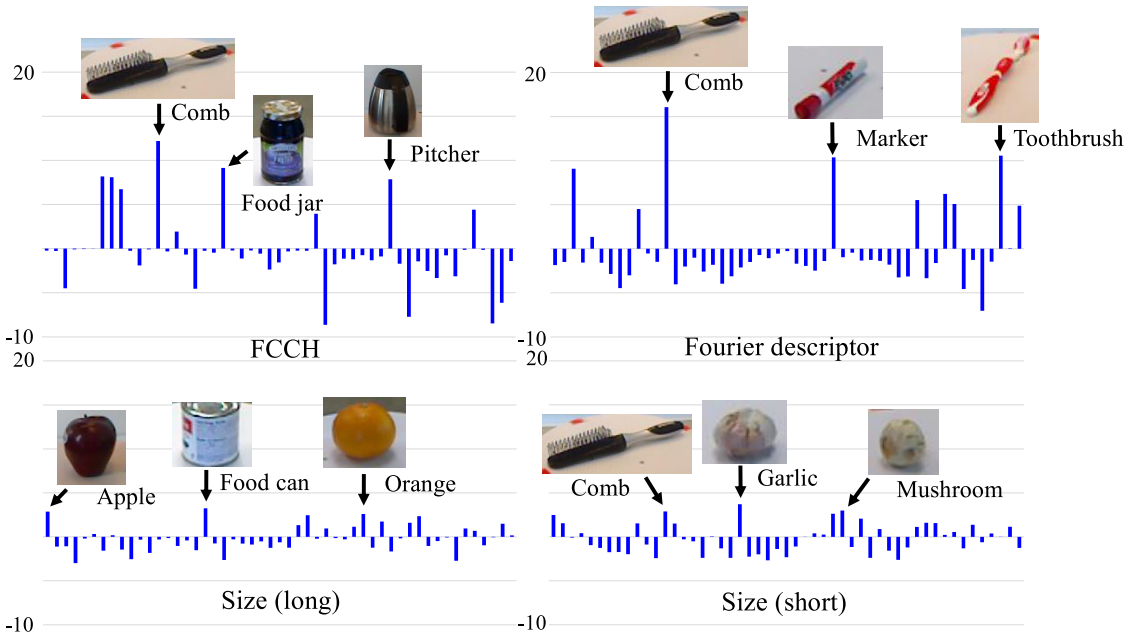

図 11 学習された Comb のカテゴリ帰属度の重み
Fig. 11 Learned weight for "Comb" with all categories.

機能を模しており，実際に学習された CNN の畳み込み層から得られるフィルタ機能と，哺乳類の初期視覚野（V1）との類似性も指摘されている [27]．そのため，本研究の学習を通して，CNN 内部にも何かしらの概念が生成されている可能性がある．図 10 より今回の実験では，AlexNet とファインチューニングした AlexNet の入力データに対する上位の識別結果を見ると，1 位にあげたカテゴリはすべて正解のカテゴリであり，その識別確率は 99.9%である．一方で，2 位と 3 位にあげたカテゴリの識別確率は，大きいものでも 10^{-3} のオーダである．このように，1 位とそれ以降の識別確率に大きな差があるため，今回の識別結果から本実験で CNN 内部に生成された物体概念について考察することができない．これについては今後の検討が別途必要であると思われる．

以上の結果から，提案手法では色，形状，大きさの 3 つの特徴を用いた物体概念に基づいて識別を行うため，そのカテゴリに識別された機序が明らかであり，CNN を用いた手法と比較して可読性が確認できる．

5.3 学習により得られた重みの評価

最後に，提案手法で学習された Comb カテゴリの重み $\boldsymbol{\theta}$ を**図 11** に示す．重み $\boldsymbol{\theta}$ は，FCCH（色），P 型フーリエ記述子（輪郭線），長手方向の大きさと短手方向の大きさの 4 つの特徴から構成される物体概念を表しており，どのカテゴリに大きな重みを持っているか調べることで，Comb カテゴリがどのような物体概念を持つか推察できる．

Comb カテゴリの重みを見ると，FCCH では，Food jar や Pitcher のような，黒，白，灰色を含むカテゴリに大きな値を持ち，P 型フーリエ記述子では，Marker や Toothbrush のような細長い形状をしたカテゴリに大きな値を持っている．長手方向の大きさでは Apple や Orange が，短手方向の大きさでは，Garlic や Mushroom に大きな値を持っているが，FCCH や P 型フーリエ記述子と比較して，重みの値が小さい．これは，大きさの特徴が Comb カテゴリの識別において，重要ではないことを表している．これらの結果から，Comb カテゴリは，色や輪郭線を重要と見なした概念を持っており，その概念は，黒，白，灰色を含んだ細長い形状をした物体であることが推察できる．これは Comb の学習画像が持つ視覚的特徴を大まかにとらえている．

表 3 学習時間
Table 3 Training time.

	Number of image data for each category	
	10 pcs	100 pcs
Proposal method	5.7 s	875 s (15 min.)
AlexNet	3,218 s (54 min.)	32,554 s (9.0 hr.)
AlexNet (Fine-tuning)	4,188 s (70 min.)	31,911 s (8.9 hr.)

5.4 処理時間による提案手法と CNN による手法の比較

表 3 に各カテゴリ 10 枚と 100 枚の学習画像を与えた場合の，提案手法と CNN を用いた手法の学習に要した時間を示す．提案手法の学習時間は，4.2 節に記述するように，カテゴリ帰属度の計算とロジスティック回帰による重みの学習までを計測範囲としている．CNN の学習時間は，4.4 節に記述するように，最大エポック数である 1,960 回目までの繰返し計算による学習を計測範囲としている．計算処理は，Ubuntu14.04，CPU Intel Core i5-2500K（3.30 GHz），RAM 8 GB の環境で行う．CNN の収束計算は，上記の環境に加えて，GeForce GTX 1080 を利用する．処理時間を

見ると，提案手法の学習時間は，各カテゴリ 10 枚で CNN の学習時間の 1/500 以下，100 枚の場合で CNN の学習時間の 1/30 以下である．この結果から，提案手法は高速に学習が行えており，学習に要する時間の観点からは，提案手法は CNN を用いた手法に対して優位性があるといえる．

6. 結論と今後の課題

本稿では，従来の物体認識手法とは異なり，人間のように曖昧かつ複数の特徴に基づいた概念を生成し，それを利用することで，物体の識別を行う方法を提案した．また，提案手法の評価により，比較的少ない学習枚数と短い学習時間で高い識別率を達成していることを確認した．さらに，物体概念（色，輪郭，大きさの各特徴の重み）の学習結果が，本稿の 1 章で定義した評価基準を満たしていることから，物体概念が確かに生成されていることが確認できた．

提案手法は CNN を用いた手法と比べて識別率の観点から優位ではないが，本稿の実験ではそれに匹敵するような高い識別率が得られた．このことは色，輪郭，大きさが物体の概念を形成するために適切な特徴であることを改めて示しているとともに，提案手法が可読性（特徴を用いて物体識別の機序の説明ができること）を有することも示している．

今後の課題として，データセット以外の認識対象に対して本手法がどの程度有効であるかを確認することがあげられる．また今回のデータセットの画像は物体が背景から切り出されているため，領域分割の手法と組み合わせ，任意の背景に物体が設置されたリアルワールドにおいて物体を識別することが将来的な課題である．

謝辞 本研究の一部は，平成 28 年度関西大学教育研究高度化促進費において，研究課題「ロボット競技会をモチベーションとしたソフトウェアに力点をおいたメカトロニクス教育」として研究費を受けた．

参考文献

[1] Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol.60, No.2, pp.91–110 (2004).
[2] Dalal, N. and Triggs, B.: Histograms of Oriented Gradients for Human Detection, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp.886–893 (2005).
[3] Freund, Y. and Schapire, R.E.: Experiments with a New Boosting Algorithm, *Proc. 13h International Conference on Machine Learning*, pp.148–156 (1996).
[4] Cortes, C. and Vapnik, V.: Support-vector Net-works, *Machine Learning*, Vol.20, No.3, pp.273–297 (1995).
[5] Vondrick, C. et al.: HOGgles: Visualizing Object Detection Features, *Proc. International Conference on Computer Vision*, pp.1–8 (2013).
[6] 秋月秀一，橋本 学：特徴的 3-D ベクトルペアを用いたばら積み部品の高速位置姿勢認識，電気学会論文誌 C，Vol.133, No.9, pp.1853–1854 (2013).
[7] Krizhevsky, A., Sutskever, I. and Hinton, G.E.: Imagenet classification with deep convolutional neural networks, *Proc. Neural Information and Processing Systems*, pp.1097–1105 (2012).
[8] Sachin, R. and Hugo L.: Optimization as a Model for Few-Shot Learning, *International Conference on Learning Representations* (2017).
[9] Nguyen, A., Yosinski, J. and Clune, J.: Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp.427–436 (2015).
[10] Ullman, S.: *High-level Vision*, The MIT Press (1995).
[11] 横澤一彦：高次視覚の認知心理学と脳機能計測，認知神経科学，Vol.1, No.2, pp.152–156 (1999).
[12] Muller, H.J., Heller, D. and Ziegler, J.: Visual search for singleton feature targets within and across feature dimensions, *Perception & Psychophysics*, Vol.57, No.1, pp.1–17 (1995).
[13] Lai, K. et al.: A Large-Scale Hierarchical Multi-View RGB-D Object Dataset, *Proc. IEEE International Conference on Robotics and Automation*, pp.1817–1824 (2011).
[14] Han, J. and Ma, K.K.: Fuzzy color histogram and its use in color image retrieval, *IEEE Trans. Image Processing*, Vol.11, No.8, pp.944–952 (2002).
[15] 上坂吉則：開曲線にも適用できる新しいフーリエ記述子，電子情報通信学会論文誌 A，Vol.67, No.3, pp.166–173 (1984).
[16] 内川恵二：表面色のカテゴリカル知覚，光学，Vol.17, No.12, pp.47–55 (1988).
[17] Bezdek, J.C., Ehrlich, R. and Full, W.: FCM: The fuzzy c-means clustering algorithm, *Computers & Geosciences*, Vol.10, No.2-3, pp.191–203 (1984).
[18] 奥富正敏：ディジタル画像処理，p.62, CG-ARTS 協会 (2004).
[19] Sharma, G., Wu, W. and Dalal, N.: The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations, *Color Research & Application*, Vol.30, No.1, pp.21–30 (2004).
[20] Frey, B.J. and Dueck, D.: Clustering by passing messages between data points, *Science*, Vol.315, No.5814, pp.972–976 (2007).
[21] Bishop, C.M.: *Pattern Recognition and Machine Learning*, pp.209–210, Springer (2006).
[22] Tang, S. et al.: Histogram of oriented normal vectors for object recognition with a depth sensor, *Proc. Asian Conference on Computer Vision*, pp.525–538 (2013).
[23] Jeff, D. et al.: DeCAF: A deep convolutional activation feature for generic visual recognition, *International Conference on Machine Learning*, pp.647–655 (2014).
[24] Jia, Y. et al.: Caffe: Convolutional Architecture for Fast Feature Embedding, *Proc. 22nd ACM International Conference on Multimedia*, pp.675–678 (2014).
[25] Deng, J. et al.: ImageNet: A Large-Scale Hierarchical Image Database, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp.248–255 (2009).
[26] Agrawal P. et al.: Analyzing the performance of multilayer neural networks for object recognition, *European Conference on Computer Vision*, pp.329–344 (2014).
[27] 岡谷貴之：画像認識のための深層学習（〈連載解説〉Deep Learning（深層学習）[第 4 回]），人工知能学会誌，Vol.28, No.6, pp.962–974 (2013).

秋本 翔平 （正会員）

2016 年関西大学システム理工学部機械工学科卒業，2018 年同大学大学院理工学研究科システム理工学専攻機械工学分野修了，現在に至る．

高橋 智一

2010 年東北大学大学院工学研究科ナノメカニクス専攻博士課程後期修了．同年関西大学システム理工学部助教，2015 年准教授，現在に至る．博士（工学）．フレキシブル吸着グリッパ，MEMS 技術を用いた小型発電デバイス，ソフトアクチュエータ等の研究に従事．IEEE，日本機械学会，電気学会各会員．

鈴木 昌人

2007 年広島大学大学院先端物質科学研究科博士課程後期修了．同年同大学ナノデバイス・システム研究センター博士研究員．2008 年より関西大学システム理工学部助教，2013 年准教授，現在に至る．博士（工学）．マイクロマシニング技術，マイクロセンサ等の研究に従事．IEEE，日本機械学会，電気学会各会員．

青柳 誠司

1988 年東京大学大学院工学研究科精密機械工学専攻修士課程修了．同年金沢大学工学部機械システム工学科助手，1995 年同大学助教授，1996 年関西大学工学部管理工学科助教授，2002 年同大学同学部システムマネジメント工学科助教授，2003 年同大学同学部システムマネジメント工学科教授，2007 年同大学システム理工学部機械工学科教授，現在に至る．博士（工学）．ロボット用センサ・アクチュエータ，ロボット制御，マイクロマシン等の研究に従事．IEEE，精密工学会，日本機械学会，電気学会等各会員．