

物体の符号化による 画像知識を取り入れた概念ベースの構築

Developing the Concept-Base Incorporated Image Knowledge with Encoding the Object

小川真路^{1*} 芋野美紗子² 土屋誠司³ 渡部広一³

Shinji Ogawa¹ Misako Imono² Seiji Tsuchiya³ Hirokazu Watabe³

¹ 同志社大学大学院理工学研究科

¹ Graduate School of Science and Engineering, Doshisha University

² 同志社大学高等研究教育機構 (理工学部)

² Organization for Advanced Research and Education (Faculty of Science and Engineering),
Doshisha University

³ 同志社大学理工学部

³ Faculty of Science and Engineering, Doshisha University

Abstract: For make smoother communication between computer and human, computer need to retain knowledge of the human language. Concept-Base is a knowledge base that gathers such knowledge in specified forms. In existing Concept-Base, the meanings of words (concepts) are defined by sets of words having a relation from something as attributes. It is considers that features in images of the words are important knowledge of the words. So this paper proposes a method of creating the Concept-Base incorporated the image knowledge with encoding the object in images.

1 はじめに

人は、ある語から関連性のある語を連想する能力がある。この連想能力をコンピュータに実現することができれば、人が行うような言葉の意味理解を表現できると考えられる。これにより、言葉を理解し、人のように返答できる会話システムの実現に役立つと考えられる。そのためには、コンピュータが語と語の関連性に関する知識を大量に保持しておく必要があり、それらを一定形式で集約し保持した知識ベースとして概念ベース[1]がある。

概念ベースは複数の電子国語辞書から機械的に構築された知識ベースであり、様々な語（概念）がそれを特徴付ける語（属性）とその重要性を表す数値（重み）の対の集合により定義されている。概念ベースでは、ある語が示すものが現実世界においてどのような事物・事象であるかを概念化し、語の意味を定義している。ここでいう意味とは、国語辞書の語義文に示されているものだけでなく、語や文を見

たときに人間が理解する内容の事だと考える。したがって、語の意味を概念として本質的に捉えるためには、語に関する様々な知識を概念知識として保持させる必要がある。

既存の概念ベースでは語の意味を言葉のみで表現しているが、ある語からそれに関して想起するイメージ、像も重要な知識であると考えられる。たとえば、人は「りんご」について理解する際、「りんご」の像を思い浮かべて、そこから「赤い」や「丸い」ものだと理解する。そのような画像中における物体の特徴を概念知識として概念ベースに付与することで、概念は外形的な特徴に関する知識を保持することが可能となる。

そこで本稿では、語だけでなく、画像中における物体を表現する特徴（画像知識）を取り入れた概念ベースの構築手法を提案する。

2 概念ベース

1章で述べたように概念ベースにおいて概念Aは m 個の属性 a_i と重み $w_i (>0)$ の対によって定義されており、一般的には次のように表される。

*連絡先：同志社大学大学院理工学研究科
〒610-0394 京都府京田辺市多々羅都谷 1-3
E-mail: dun0140@mail4.doshisha.ac.jp

$$\text{概念}A = \{(a_1, w_1), (a_2, w_2), \dots, (a_m, w_m)\} \quad (1)$$

ここで、属性 a_i を概念 A の一次属性と呼ぶ。属性 a_i も概念ベースの中で1つの概念として定義されているため、 a_i からも同様に属性を導くことができる。 a_i の属性 a_{ij} を概念 A の二次属性と呼ぶ。概念ベースの例を図1に示す。

医者	医者, 0.36	患者, 0.21	病院, 0.10	...	一次属性
	医者, 0.36	患者, 0.21	病院, 0.10	...	
	患者, 0.21	病人, 0.23	医院, 0.25	...	二次属性
	病院, 0.10	看病, 0.11	手術, 0.18	...	
:	:	:	:	:	:

図1: 概念「医者」を二次属性まで展開した場合の例

図1に示した概念「医者」の属性「医者」、「患者」、「病院」を概念「医者」の一次属性と呼ぶ。そして属性もまた概念ベースにおいて定義されているため、これらの一次属性を概念としてみなしさらに属性を導くことができる。たとえば「病院」を概念とみなすと「病院」、「医院」、「手術」といった属性が導ける。これらを元の概念「医者」の二次属性と呼ぶ。このように、導いた属性からさらに属性を導けることから概念ベースは n 次の属性の連鎖構造であるといえる。

3 構造化概念ベース

既存の概念ベースの応用として、属性を品詞や語関係ごとに整理した概念ベース[2]（以下、構造化概念ベース）があり、約4万5千の概念が存在する。これは、属性を名詞、形容詞および動詞ごとに品詞別でグループ化し、各グループの中で概念との語関係（同義や類義など）を持たせた構造で構築された概念ベースである。構造化概念ベースの概念の定義と例を図2に示す。重みについては便宜上省略している。たとえば、概念 N の属性 n_{sa} は、概念 N と同義関係であり名詞の属性である。図2では、名詞概念の定義を示したが、形容詞概念、動詞概念も同様の定義である。

概念	属性		
N	$\{ n_{sa}, n_{si}, n_r, n_i, n_r, a_{sa}, a_{si}, a_o, a_i, a_r, v_{sa}, v_{si}, v_o, v_i, v_r \}$		
N : 名詞概念	sa : 同義語	i : 反意語	
n : 名詞属性	si : 類義語	r : 共起語	
a : 形容詞属性	t : 上位語	o : 名詞化	
v : 動詞属性			
(例)	概念「バス」 = { 乗合自動車, 二階建バス, ... 大勢人, 大型, ... }		
	名詞 同義	名詞 類義	名詞 共起

図2: 構造化概念ベースにおける概念の定義と例

4 関連度

関連度とは、概念ベースに定義されている概念間の関連性を定量的に表現した値である。関連度は0.0から1.0の間で値が変動し、概念間の関連が強いほど大きな値を示す。以下に関連度の算出に用いる一緻度および関連度計算方式[3]について述べる。

4.1 一緻度

ある概念 A, B について、その一次属性を a_i, b_j 、重みを u_i, v_j とする。それぞれが持つ属性数が L 個、 M 個 ($L \leq M$) とすると、概念 A, B はそれぞれ

$$A = \{(a_1, u_1), (a_2, u_2), \dots, (a_L, u_L)\} \quad (2)$$

$$B = \{(b_1, v_1), (b_2, v_2), \dots, (b_M, v_M)\} \quad (3)$$

と表現される。このとき概念 A と概念 B の一緻度 $DoM(A, B)$ は以下のように定義される。

$$DoM(A, B) = \sum_{a_i=b_j} \min(u_i, v_j) \quad (4)$$

$$\min(u_i, v_j) = \begin{cases} u_i & (u_i \leq v_j) \\ v_j & (u_i > v_j) \end{cases} \quad (5)$$

$a_i = b_j$ は属性同士が表記的に一致した場合を示している。つまり、一緻度とは概念 A と概念 B それぞれが持つ属性のうち、小さい方の重みを足し合わせたものとなる。

4.2 関連度計算方式

関連度を算出する概念同士の一次属性全ての組み合わせに対して一緻度計算を行い、一緻度の高い属性の組み合わせから順に対応を決定する。(2)式の概念 A の属性順を固定した上で、一緻度が最大となる組み合わせに概念 B の属性を並べ替えたものを以下のように定義する。

$$B = \{(b_{x1}, v_{x1}), (b_{x2}, v_{x2}), \dots, (b_{xM}, v_{xM})\} \quad (6)$$

これら概念 A, B についての関連度 $DoA(A, B)$ を以下の式で定義する。

$$DoA(A, B) = \sum_i DoM(a_i, b_{xi}) \times \frac{(u_i + v_{xi})}{2} \times \frac{\min(u_i, v_{xi})}{\max(u_i, v_{xi})} \quad (7)$$

5 BoF による画像表現

BoF (Bag-of-Features) [4]とは、画像中の物体を認識する物体認識の手法であり、画像を局所特徴量の集合と見なし、画像を局所特徴量の出現頻度ヒストグラムとして表現する手法である。ここで局所特徴量とは、人の顔における目や鼻といったように、ある物体を構成するパーツのことであり、画像中のそのような特徴的な部分（特徴点）を特徴ベクトルで記述したものである。本稿では局所特徴量の記述にSURF[5]を用いる。図3にBoFの例を示す。

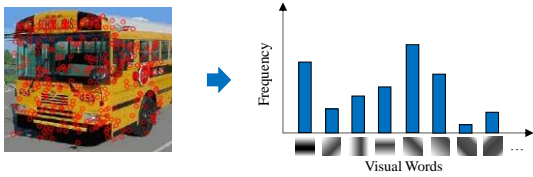


図 3: BoF による画像表現

具体的な手法として、事前に複数の画像から局所特徴量を抽出し、その抽出した全ての局所特徴量を用いてクラスタリングを行う。そして、1枚の画像から局所特徴量を抽出し、クラスタリングにより生成された各重心に属する局所特徴量の数をヒストグラムとして表現する。クラスタリングにより生成された各重心を Visual Words (以降, VW) とする。VW の数はクラスタの数に比例するため可変である。

6 画像知識を取り入れた概念ベース

本章では、画像知識を取り入れた概念ベースの構築手法について述べる。

6.1 画像の情報源

概念ベースに画像知識を取り入れる際の画像の情報源として、Caltech-256[6]を使用する。本稿では、Caltech-256 から 30 個の物体名を選出し、それぞれ 50 枚の画像を選出した。その内の 45 枚の画像を概念ベースの構築に用いて、残りの 5 枚は後述する物体認識システムの入力で用いた。また、30 物体各 50 枚の画像の内、45 枚の画像はランダムで選出し、5 枚は目視で選出した。

6.2 画像中における物体の符号化

BoF により画像中における物体の特徴 VW を取得し、それを概念とみなすことで、物体を概念集合で表現できると考えた。これにより、(1)式のように画像中の物体を 1 つの概念として捉えることができる。この処理を物体の符号化と定義する。

6.2.1 画像 1 枚を用いた物体の符号化

BoF によって抽出したそれぞれの VW に対して VW1, VW2, VW3, ... とラベル付けし、それぞれを概念と定義する。また、VW の値が大きいほど画像を強く特徴付けているという意味合いであるため、VW の値を画像に対する重要度として捉えることができる。この重要度は、2 章で述べた概念ベースの構築に当てはめると、各属性の重みと対応付けできる。したがって、図 4 で示すように、1 枚の画像中の物体を VW と重要度を表す重みの集合、つまり概念として表現することができる。ここで、重みが 0 である VW は、その画像中の物体の意味特徴を表現して

いないとして、その集合から削除する。

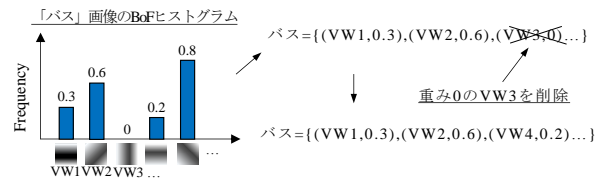


図 4: 画像 1 枚を用いた物体の符号化

6.2.2 画像複数枚を用いた物体の符号化

6.1 節で述べたように、1 つの物体名に対して複数の画像が存在する。これら複数の画像から抽出した全ての特徴を統合することで、1 つの物体が普遍的に持つ特徴を取得することができると考えた。そこで、複数の画像それぞれに対して物体の符号化を行い、各 VW に対して平均値、中央値のいずれかで統合する。

6.3 構築方法

本節では、物体の符号化により画像知識を取得し、それを取り入れた概念ベースの具体的な構築手法について述べる。本稿では、物体の符号化によって得られた概念 VW を画像概念と定義し、属性中に存在する VW を画像属性と定義する。また、画像概念および画像属性との区別のために、語で定義される概念を特に語概念と定義し、属性中に存在する語概念を語属性と定義する。

6.3.1 語概念に対する画像属性の追加

6.2 節で述べたように、物体の符号化によって取得した特徴は、ある物体についての画像中の特徴を表現している。そのため、物体の符号化によって取得した VW の集合を、画像属性として物体名に対応した語概念に追加する。追加する際は、構造化概念ベースの概念の定義において、語属性の末尾に画像属性として追加する。これにより、語概念に画像知識を付与することができる。追加した画像属性に対する重み付け手法として言語処理の分野で用いられる *tf-idf* [7] を用いる。その式を以下に示す。

$$w(d, t) = tf(d, t) \cdot \left(\log \frac{N}{df(t)} + 1 \right) \quad (8)$$

$w(d, t)$ は記事 d における索引語 t の重要度、 $tf(d, t)$ は記事 d における索引語 t の出現頻度、 N は記事数、 $df(t)$ は索引語 t が出現する記事の数である。画像属性の重み付けでは、各種 30 物体に対して物体の符号化を行い、取得した VW と重みの集合で表現された空間を用いる。

6.3.2 画像概念に対する画像属性の追加

物体の符号化で得られた VW を画像概念として登録し、属性を追加する。画像概念に対する画像属性の追加では、物体の符号化により取得した複数の物

体名と VW の集合で表現された空間を用いる。ある画像概念に対する画像属性として、ある物体名の中で画像概念 VW と共に出現する VW を取得する。これをすべての物体名に対して行う。図 5 に画像概念に対する画像属性の取得の例を示す。

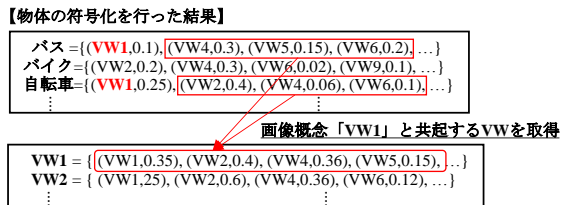


図 5: 画像概念に対する画像属性の取得の例

たとえば、画像概念 VW1 の画像属性として、VW1 を属性として持つ「バス」や「自転車」の VW、つまり VW4 や VW2 を取得する。画像属性 VW4 や VW2 は、一枚の画像を形成する際に画像概念 VW1 と関わっていると考え、それらを共起属性として取得した。ここで、共起関係により画像属性を取得すると、同じ表記の VW が重複して取得される。重複した VW は、お互いの重みを加算し統合する。追加した画像属性の重み付け手法として前項と同様に *tf-idf* を用いる。重み付けの際は、画像概念に対する画像属性と重みの集合で表現された空間で行う。

6.3.3 画像概念に対する語属性の追加

ある画像概念に対する語属性として、画像概念 VW を属性として持つ物体名を取得する。語属性の取得の例を図 6 に示す。

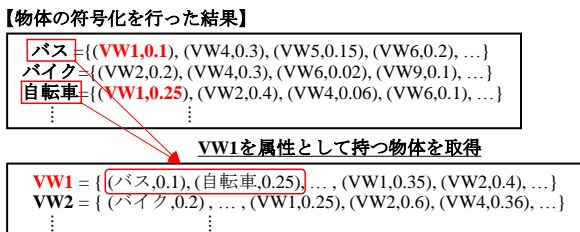


図 6: 画像概念に対する語属性の取得の例

たとえば、画像概念 VW1 の語属性として、VW1 を属性として持つ「バス」や「自転車」を取得する。VW1 は「バス」や「自転車」の物体を表現している一要素であるため、VW1 にとってそれらと関連性が強いと考えた。追加した語属性の重みとして、6.3.1 項で語概念に対して画像属性に付与された重みを、VW に対するその物体の重みとして捉え、画像概念に対する語属性の重みとして付与する。

6.3.4 構築結果

VW の総数を 500 に設定し、物体の符号化において平均値、中央値での統合それぞれに対して概念ベースを構築した。平均値で統合化した概念ベースに

おける語属性のみの平均属性数は 24.4 個、画像属性のみのみは 499.5 個、それらを合わせた平均属性数は 523.9 個となった。一方、中央値で統合化した際は、それぞれ 24.4 個、103.7 個、128.1 個となった。図 7 に構築した概念ベースの例を示す。

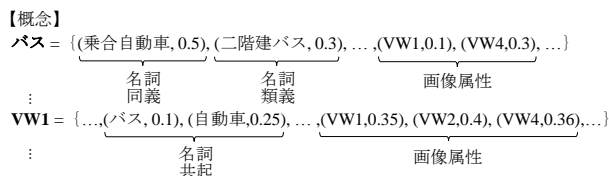


図 7: 画像知識を取り入れた概念ベースの例

7 評価

画像知識を取り入れた概念ベースに対する評価として、物体認識により評価を行う。

7.1 評価方法

概念ベースを用いた物体認識システムを構築し、それにより概念ベースの評価を行う。精度比較の対象として、物体認識における既存手法であるヒストグラムインターセクション[8]と SVM[9]を用いる。本評価では目視により選出した 30 物体各 5 枚の計 150 枚の画像を物体認識システムに入力し、出力された物体名が入力画像中の物体名と一致する時に正解として、その正解率を精度とする。

7.1.1 概念ベースを用いた物体認識

画像知識を取り入れた概念ベースを用いた物体認識システムを図 8 に示す。

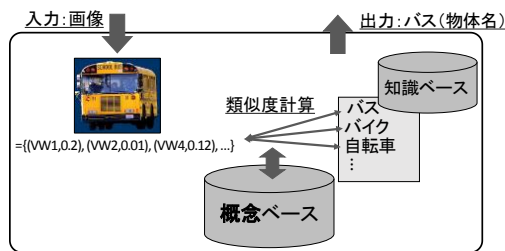


図 8: 概念ベースを用いた物体認識システム

本システムの入力は 1 枚の画像であり、出力は物体名である。知識ベースには複数の物体名が格納されており、本稿では 30 個の物体名が存在する。入力画像に関して物体の符号化により取得した VW の集合を一次属性とみなすことで、1 枚の画像を概念とみなすことができる。これにより、概念ベースを用いて入力画像と知識ベース内の物体名について関連度または一致度計算により類似度を算出できる。知識ベース内の全ての物体名と入力画像との類似度を計算し、最大値となった物体名を入力画像中に写る物体名と判断し出力する。

関連度の計算では、4章の関連度計算方式と同様の手順で行うが、属性を二次展開して一致度計算を行う際、語属性と画像属性の数を考慮して一致度を算出する。これにより、語と画像を考慮した計算を行うことができる。具体的に、単に属性を重み降順で取得し一致度の計算を行うのではなく、語属性と画像属性それぞれを使用属性数の半数ずつ重み降順で取得し、その属性を用いて一致度を算出する。

7.1.2 ヒストグラムインターセクション

既存研究[8]では、ヒストグラムインターセクション（以降、HI）を用いることで物体同士の類似度を計算し、物体認識を実現している。HIとは、2つのヒストグラム間の類似度を計算する方法である。以下にその式を示す。

$$H = \sum_i \min(H_n(i), H_m(i)) \quad (9)$$

$H_n(i)$ はヒストグラムの*i*番目の値であり、 $\min()$ は2つのヒストグラムの値を比較し、小さい方の値を求める。ヒストグラムの全ての要素を合計1.0になるように正規化されているため、類似度の値は0.0から1.0の実数値となる。入力画像と30物体各45枚の計1350枚の全画像とのHIを求め、最大値となった物体名を出力する。

7.1.3 Support Vector Machine (SVM)

SVM[9]は、2値分類問題を解くために考えられた分類器であり、パターン認識の手法である。本稿では複数クラスに対応したLIBSVM[10]を用いる。本稿におけるSVMを用いた物体認識では、30物体各45枚の計1350枚の画像を用いて学習を行う。そして、入力画像中の物体を識別し、その物体名を出力する。本稿では、RBFカーネル[9]を用いた。

7.2 評価結果

表1に評価結果を示す。物体の符号化の際に平均値で統合化した概念ベースおよび中央値で統合化した概念ベースを比較している。DoMは一致度計算を用いた手法である。DoA(I)は関連度を算出の際に画像属性のみを用いた手法であり、DoA(I+W)は画像属性と語属性を用いた手法である。表1における概念ベースを用いた評価結果では、最大精度となった使用属性数における結果を示す。

表1: 精度評価比較 (%)

	平均値			中央値			HI	SVM
	DoM	DoA(I)	DoA(I+W)	DoM	DoA(I)	DoA(I+W)		
	45.3	34.7	34.7	32.7	24.0	26.0	30.0	41.3

表1の結果から、平均値で統合化した概念ベースを用いた一致度計算手法がSVMと同程度であるが

最も高い精度となった。また、関連度計算方式よりも一致度計算を用いた手法の方が高い精度となることがわかる。さらに、関連度計算方式において、画像属性だけでなく語属性も用いることで、画像属性のみを用いた時よりも一部で良い結果が得られている。

次に精度の高い平均値で統合化した概念ベースにおけるDoMとSVM、HIの3つの手法に対して、VWの総数の変化による精度評価を行った。その結果を図9に示す。

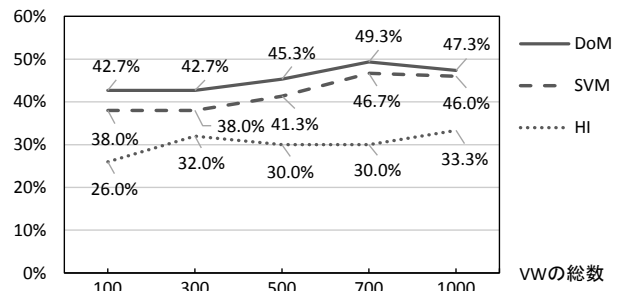


図9: VWの総数の変化による精度評価

結果から、DoMはHIよりも高い精度となった。また、VWの総数が大きい時、DoMとSVMは同程度の精度となり、VWの総数が小さくなるに従い、DoMとSVMの精度の差が大きくなった。

8 考察

画像知識を取り入れた概念ベースに対して、物体認識による評価結果についての考察を述べる。

8.1 使用属性数の考慮について

平均値で統合化した概念ベースを用いた際、使用属性数を変化させた時の評価結果を図10に示す。

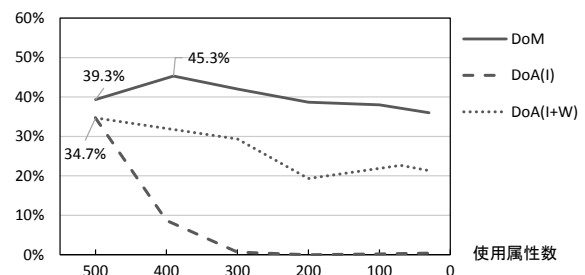


図10: 使用属性数の変化による評価結果

一致度計算において使用属性数を考慮することで精度が向上した。これは、重み上位の属性を一致度計算に用いることで、重みの低い属性を雑音と見なして削除することができ、雑音の影響を抑制できたためだと考えられる。このように、使用属性数を考

慮することで、雑音となる属性を削除することができ、精度の向上を示したと考えられる。ただし、関連度計算方式を用いた手法では、低い精度となっている。これは、二次属性を展開して関連度を計算しており、画像概念が持つ属性として相応しくない属性が多く付与されているためであると考えられる。したがって、概念ベースの精練（不適切な属性の削除および新たな重み付け）が必要となる。

8.2 VW の総数および画像の枚数について

図9の結果から、VWの総数が大きい時、3つの手法においてVWの総数が小さい時よりも高い精度を得られている。VWの総数を大きく設定することは、物体を表現できる特徴の幅（数）が広がることを意味しており、それにより物体をVWにより細かく表現することができたと考えられる。一方、VWの総数が小さいと、物体を表現できる特徴の幅が狭くなり、物体をVWにより上手く表現することが困難になる。図9の結果ではVWの総数が小さくなるに従い、DoMとSVMの精度の差が大きくなる傾向を示している。DoMの手法では、物体の符号化の際、統合化により複数の画像から物体の代表的な特徴を取得し、さらに使用属性数を考慮することで雑音を除去している。これにより、物体を特徴付けることが困難な状況であっても、上手く物体を表現する特徴を取得できたと考えられる。

物体を特徴付けるための情報が少ない時でもDoMによる手法では上手く物体を表現できることを検証するために、使用する画像の枚数を極端に少ない状況で評価比較を行った。そこで、30物体に対して各10枚の画像を選出し、その内の5枚をランダムに選出した。その画像を概念ベース構築やSVMの学習画像として使用し、残り5枚を入力画像とした。評価結果を図11に示す。

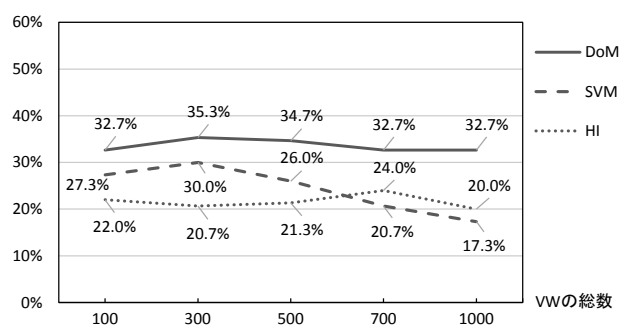


図11: 画像枚数が少ない状況での評価結果

結果からDoMがSVM, HIより高い精度となった。したがって、使用する画像の枚数が少ない時やVWの総数が小さいといった、物体の代表的な特徴を表

現する上で厳しい環境であっても、本手法により概念ベースにおいて物体の特徴を上手く属性として表現することができたと考えられる。

9 むすび

本稿では、語概念のみで定義された既存の概念ベースを用いて、画像知識を取り入れた概念ベースの構築を行った。本稿の手法により構築した概念ベースでは、物体を表現する代表的な特徴を取得する上で厳しい状況であっても、物体の特徴を上手く属性として表現することができたとと言える。今後は、画像概念が持つ属性に問題があると考えられるため、概念ベースに対して精練を行う必要がある。

謝辞

本研究の一部は、科学研究費補助金（若手研究（B）24700215）の補助を受けて行った。

参考文献

- [1] 笠原要, 松澤和光, 石川勉: 国語辞書を利用した日常語の類似性判別, 情報処理学会論文誌, Vol. 38, No. 7, pp. 1272-1283, (1997)
- [2] 小川真路, 芋野美紗子, 土屋誠司, 渡部広一: 概念の多義性を考慮した属性構造化による概念ベースの構築, FIT2013, E-019, pp. 223-224, (2013)
- [3] 荒木孝允, 奥村紀之, 渡部広一, 河岡司: 比較対象概念の共通属性を重視する動的関連度計算方式, 同志社大学理工学研究報告, Vol. 48, No. 3, pp. 14-24, (2007)
- [4] G. Csurka, C.R. Dance, L. Fan, C. Bray: Visual Categorization with Bags of Keypoints, European Conference on Computer Vision, pp. 1-22, (2004)
- [5] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool: SURF: Speeded Up Robust Features, Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346-359, (2008)
- [6] Caltech 256 image dataset, http://www.vision.caltech.edu/Image_datasets/Caltech256/
- [7] 徳永健伸 (編): 情報検索と言語処理, 東京大学出版会, (1999)
- [8] 古川拓也, 吉村枝里子, 土屋誠司, 渡部広一: SURF特徴量を用いたBoF法による物体認識, 研究報告知能システム(ICS), 2014-ICS-174(5), pp. 1-8, (2014)
- [9] 大北剛 (訳): サポートベクターマシン入門, 共立出版 (株), (2005)
- [10] Chang, C. C. and Lin, C. J.: LIBSVM: A Library for Support Vector Machines, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>