2D3-OS-12b-1

# 混合マルチモーダルHDPによる多様な概念の形成

Various Concept Formation Using an Infinite Mixture of Multimodal HDP

中村友昭 Tomoaki Nakamura　安藤義記 Yoshiki Ando　長井隆行 Takayuki Nagai

電気通信大学大学院情報理工学研究科
Faculty of Informatics and Engineering, The University of Electro-Communications

We propose a method for a robot to form various concepts. The robot uses its embodiment to obtain visual, auditory, and haptic information by grasping, shaking, and observing objects. At the same time, a user teaches the robot object features through speech. From these kinds of information, the robot can form object concepts. The information obtained by the robot is converted into Bag-of-Words representations, which are classified into categories. The proposed method is based on a stochastic model, and objects can be classified by estimating their parameters. We introduce the Chinese restaurant process into the multimodal hierarchical Dirichlet process. This model is an infinite mixture of models and enables the robot to form various concepts such as object type, color, and so on. Therefore, the robot can form not only object concepts but also various concepts that are not represented by object concepts (e.g., color). Also, because the proposed method is based on a stochastic model, it makes it possible for the robot to estimate category and unobserved information of unseen objects. We implement the proposed method on a robot and show that it can form various concepts and perform various estimations for unseen objects.

## 1. はじめに

事物のカテゴリ分類は，人間の認知機能において重要な役割を果たしていることが指摘されている．人間はカテゴリを形成することで，経験した物事を全て参照することなく，必要最小限の認知的処理によってより多くの情報を得ることができる．また，カテゴリ分類の重要性は，経験を通して形成したカテゴリを利用した予測が可能な点にある．人間は，未知の物事に対しても様々な予測を行い，柔軟に対応している．さらに自己組織的に分類されたカテゴリを「概念」と捉えると，概念と単語との結びつきから，カテゴリを通した予測が言葉の意味理解につながっていると考えることができる．このように，カテゴリ分類とカテゴリを通した予測は人間の柔軟な知能の基盤であり，知能ロボットにもこのような能力を持たせることが必要であると考える．

これまで我々は，自然言語処理の分野で盛んに研究されてきた統計モデルの一つである Latent Dirichlet Allocation (LDA) をベースに，物体カテゴリを教師なしで形成する手法を提案してきた [Nakamura 12]．これらの研究では，物体から取得できる視覚や聴覚，触覚などのマルチモーダル情報を LDA によりカテゴリ分類することで，ロボットがペットボトルや瓶，ぬいぐるみといった人間の感覚に即した物体のカテゴリ (概念) を形成できることを示した．しかしながら，これらの研究は物体のカテゴリの形成にのみ注目しており，形成されるカテゴリは単一の構造である．その結果，色に注目した概念，音に注目した概念など，様々な視点での概念が形成されないため，物体名などの名詞は表現できるものの，形容詞や副詞などを表現することができない．

そこで本稿では，これまで物体概念の構築に用いてきたノンパラメトリックベイズモデルである Hierarchical Dirichlet Process を，マルチモーダルカテゴリゼーションへと拡張した Multimodal Hierarchical Dirichlet Process (MHDP) [中村 13] に，Chinese Restaurant Process (CRP) [Aldous 85] を導入し，各モダリティの情報の比率が異なる複数の概念を形成することで，それぞれのモダリティの情報により重みづけされた概念を形成する手法を提案する．提案手法により，ロボットは物体から得た情報を通じて物体の概念を形成するだけではなく，物体概念では表現できない色を表現する概念など，複数の概念を獲得することが可能となる．また，形成した概念を利用することで，未知物体に対するカテゴリの予測や，未観測情報の予測を行うことが可能となる．

連絡先: 中村友昭，電気通信大学大学院情報理工学研究科，東京都調布市調布ヶ丘 1-5-1，naka_t@apple.ee.uec.ac.jp

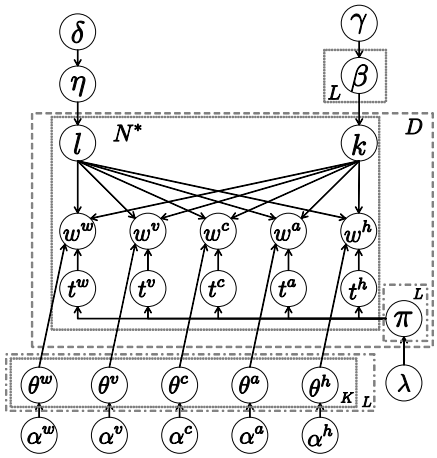

図 1: MHDP を概念形成に用いた IMoMs

## 2. 提案手法

提案手法は，図 1 に示すグラフィカルモデルに基づいており，物体のカテゴリ分類は，そのパラメータを推定する問題となる．本稿では，図 2 に示すロボットを用い，自律的に物体の視覚・聴覚・触覚情報などのマルチモーダル情報を取得する [Araki 12]．本稿では，これらマルチモーダル情報と，人により教示された単語情報を入力データとして用いる．

### 2.1 マルチモーダル情報の処理

ここでは取得するマルチモーダル情報と，その処理に関して述べる．

**視覚情報**

まず観測した物体の画像を複数枚取得する (後述する実験では，各物体に対して 36 枚の画像を取得した)．本稿では特徴量

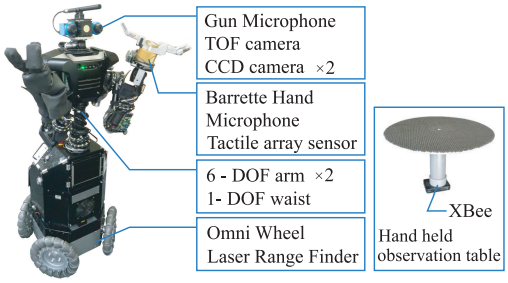


図 2: ロボットプラットフォーム

として 128 次元の DSIFT を用い，これにより 1 枚の画像から多数の特徴ベクトルを得ることができる [Lowe 04]．これらの特徴ベクトルを，学習画像とは関係のない背景画像から計算した 500 の代表ベクトルを用いてベクトル量子化することで得られる 500 次元のヒストグラムを視覚情報として取り扱う．

さらに，2 つ目の視覚情報として，Lab 表色系の補色次元 $a$ 及び $b$ の 2 次元ヒストグラムを用いた．ビンの数はそれぞれ 5 とし，合計 25 次元のヒストグラムとした．

**聴覚情報**

ロボットが物体を振ることで発生する音を聴覚情報に利用する．取得した音情報は 0.2[sec] 毎のフレームに分割し，フレーム毎の特徴量に変換する．特徴量としては，音声認識でよく利用されている MFCC を用い，各フレームは 13 次元の特徴ベクトルとなる．これにより，物体から発生した音から複数の特徴ベクトルを得ることができる．この特徴ベクトルを，あらかじめ計算した 50 個の代表ベクトルによりベクトル量子化を行い，各代表ベクトルの発生頻度を表すヒストグラムを聴覚情報として使用する．

**触覚情報**

触覚情報には，ロボットが実際に物体を掴むことで，触覚センサにより得られる時系列データを用いる．取得したデータは近似を行い，近似パラメータを各センサの特徴ベクトルとして扱う [中村 10]．さらに k 平均法により予め計算した 15 の代表ベクトルを用いてベクトル量子化を行い，15 次元ヒストグラムを触覚情報として用いる．

**単語情報**

ロボットは，物体を観察中に人から発せられた教示発話を単語情報として利用する．教示発話のうち，物体の特徴を表す単語を選び，最終的に，単語の発生頻度ヒストグラムを単語情報として用いる．

## 2.2 Infinite Mixture of Models

LDA などの文書のトピックモデルに CRP を導入することで，複数のトピックモデルの混合によりデータを表現することを可能としたのが，無限混合モデル (Infinite Mixture of Models:IMoMs) である．IMoMs では，共起する特徴量同士が一つのトピックモデルに集中することで，特徴量で重みづけされた複数のトピックモデルを同時に学習することを可能とした枠組みである．図 1 が MHDP を用いた無限混合マルチモーダル階層ディリクレ過程 (IMoMHDPs) のグラフィカルモデルである．このモデルでは物体や色を表すモデルの混合によって物体を表現するモデルとなっている．図 1 において，$\eta$ は MHDP モデルが生成される確率を表す多項分布のパラメータであり，$l$ は生成された MHDP のモデルのインデックスである．さらに，$k$ は物体のカテゴリであり，$\beta$ をパラメータとする多項分布から生成される．また，$\beta$ は $\gamma$ をパラメータとする CRP によって生成される多項分布のパラメータである．$t^*$ は物体の特徴量に割り当てられるクラスを表しており，$\pi$ をパラメータとする多項分布から生成される．また，$\pi$ は $\lambda$ をパラメータとする CRP によって生成される多項分布のパラメータである．$w^v$, $w^c$, $w^a$, $w^h$, $w^w$ は，それぞれ視覚・色・聴覚・触覚・単語情報であり，$\theta^*$ をパラメータとする多項分布から生成される．また，$\theta^*$ は，$\alpha^*$ をパラメータとするディリクレ事前分布によって決定される．

## 2.3 IMoMs によるカテゴリ分類

物体の分類は図 1 のモデルのパラメータを物体の情報を用いて学習することに相当する．本稿では，モデルのパラメータ学習に Gibbs Sampling を用いる．MHDP を概念形成に用いた IMoMs では，物体 $d$ のモダリティ $m$ の $n$ 番目の単語へ割り当てられるモデル $l^m_{dn}$，クラス $t^m_{dn}$，及びカテゴリ $k^m_{dn}$ を事後分布からサンプリングすることで，パラメータの推定を行う．

### 2.3.1 モデルのサンプリング

$d$ 番目の物体のモダリティ $m$ の $n$ 番目の特徴量へ割り当てられるモデル $l^m_{dn}$ を以下の式からサンプリングを行う．

$$P(l^m_{dn}|w^m_{dn}, \mathbf{W}^{lm}, \delta) \propto P(l^m_{dn}|\delta)P(w^m_{dn}|\mathbf{W}^{lm}) \tag{1}$$

式 (1) の 1 つ目の項は，CRP によって生成される多項分布であり，モデルが $l$ となる確率は以下のようになる．

$$P(l^m_{dn} = l|\delta) = \begin{cases} \frac{N_l}{\delta+N-1} & (l = 1, \cdots, L) \\ \frac{\delta}{\delta+N-1} & (l = L+1) \end{cases} \tag{2}$$

ここで，$N_l$ はモデル $l$ に割り当てられている特徴量の数を表しており，$N$ は特徴量の総数を表している．また，$L$ はモデル (MHDP) の総数である．$l = L+1$ が選択されることは，新たなモデルが生成されることを意味する．CRP を事前分布として用いることにより，多くの特徴量を表現しているモデルが選択されやすくなる．式 (1) の 2 つ目の項は，MHDP より特徴量が生成される確率である．

### 2.3.2 カテゴリのサンプリング

次に，全ての物体の全ての特徴量に対してモデルが割り当てられているとして，カテゴリのサンプリングを行う．まず，$d$ 番目の物体のモダリティ $m$ の $n$ 番目の特徴量へ割り当てられるクラス $t^{lm}_{dn}$ を以下の式よりサンプリングする．

$$\begin{aligned} t^{lm}_{dn} &\sim P(t^{lm}_{dn}|\mathbf{W}^{lm}_{-dn}, \lambda) \propto P(w^m_{dn}|\mathbf{W}^{lm}_{k=k^l_{dt}})P(t^{lm}_{dn}|\lambda) \\ &= \begin{cases} P(w^m_{dn}|\mathbf{W}^{lm}_{k=k^l_{dt}})\frac{N^l_{dt}}{\lambda+N^l_d-1} & (t = 1, \cdots, T^l_d) \\ P(w^m_{dn}|\mathbf{W}^{lm}_{k=k^l_{dt}})\frac{\lambda}{\lambda+N^l_d-1} & (t = T^l_d+1) \end{cases} \end{aligned} \tag{3}$$

ただし，$T^l_d$ はモデル $l$ において物体 $d$ に割り当てられているクラスの数であり，$N^l_{dt}$ はモデル $l$ において，物体 $d$ の特徴量のうち，クラス $t$ が割り当てられている特徴量の総数である．また，$N^l_d$ はモデル $l$ における物体 $d$ の特徴量の数である．さらに，$\mathbf{W}^{lm}_{-dn}$ はモデル $l$ における $d$ 番目の物体のモダリティ $m$ の $n$ 番目の特徴量を除いた特徴量の集合を表現している．$P(w^m_{dn}|\mathbf{W}^{lm}_{k=k^l_{dt}})$ は，特徴量 $w^m_{dn}$ がカテゴリが $k^l$ から生成される確率を表しており，以下の式より求めることができる．

$$P(w^m_{dn}|\mathbf{W}^l_k) = \frac{N^{lm}_{kw^m_{dn}} + \alpha^{lm}}{N^{lm}_k + d^m\alpha^m} \tag{4}$$

ただし，$N^{lm}_k$ は，モデル $l$ において，カテゴリ $k^l$ であるモダリティ $m$ の特徴量の数を表しており，$N^{lm}_{kw^m_{dn}}$ はモデル $l$ において，$w^m_{dn}$ と同じ特徴量に，カテゴリが $k^l$ が割り当てられた数を表している．また，$d^m$ はモダリティ $m$ の次元数である．

次に，クラス $t^l$ に割り当てられているカテゴリ $k^l_{dt}$ は，クラス $t^l$ に割り当てられている特徴量 $\mathbf{W}^l_{dt}$ を除いた特徴量の集合 $\mathbf{W}^l_{-dt}$ を条件とした条件付確率からサンプリングする．

$$\begin{aligned} k^l_{dt} &\sim P(k^l_{dt}|\mathbf{W}^l_{-dt}, \gamma) = P(\mathbf{W}^l_{dt}|\mathbf{W}^l_k)P(k^l_{dt}|\gamma) \\ &= \begin{cases} P(\mathbf{W}^l_{dt}|\mathbf{W}^l_k)\frac{C^l_k}{\gamma+C^l-1} & (k = 1, \cdots, K^l) \\ P(\mathbf{W}^l_{dt}|\mathbf{W}^l_k)\frac{\gamma^l}{\gamma+C^l-1} & (k = K^l) \end{cases} \end{aligned} \tag{5}$$

ただし，$K^l$ はモデル $l$ におけるカテゴリの数を表しており，$C^l_k$ はモデル $l$ においてカテゴリ $k^l$ に属するクラスの数を表している．また，$C^l$ はモデル $l$ におけるクラスの数である．さらに，$P(\mathbf{W}^l_{dt}|\mathbf{W}^l_k)$ は，モデル $l$ において，クラス $t^l$ に割り当てられている物体 $d$ の特徴量の集合 $\mathbf{W}^l_{dt}$ がカテゴリ $k^l$ から生成される確率を表しており，式 (4) の積より以下のようになる．

$$P(\mathbf{W}^l_{dt}|\mathbf{W}^l_k) = \prod_{w^m \in \mathbf{W}^l_{dt}} P(w^m|\mathbf{W}^l_k) \tag{6}$$

ただし，$\mathbf{W}^l_k$ はモデル $l$ においてカテゴリが $k^l$ である特徴量の集合である．式 (3) によるサンプリングをを全ての物体の全ての特徴量に対して繰り返し，式 (5) によるサンプリングを全ての物体の全てのクラスに対して繰り返すことで，パラメータの推定を行う．最終的に，収束した値を用いることで，モデル $l$ において，物体 $d$ がカテゴリ $k^l$ に属する確率は

$$P(k^l|\mathbf{W}^l_d) = \frac{\sum_{t^l}^{T^l_d} \delta_{k^l}(\hat{k^l_{dt}}) \hat{N^l_{dt}}}{N^l_d} \tag{7}$$

となる．ただし，$\mathbf{W}^l_d$ はモデル $l$ における $d$ 番目の物体の全ての特徴量であり，$\hat{N^l_{dt}}$，$\hat{k^l_{dt}}$ はそれぞれ推定されたパラメータ $N^l_{dt}$，$k^l_{dt}$ の値である．また，$\delta_a(b)$ は $a = b$ のとき 1，さもなくば 0 となるデルタ関数である．

### 2.4 未知物体の認識

学習した IMoMs を用いることで，未知物体のカテゴリを認識することが可能となる．未知物体 $\bar{d}$ のマルチモーダル情報 $w^v_{\bar{d}}$，$w^a_{\bar{d}}$，$w^h_{\bar{d}}$，$w^c_{\bar{d}}$，$w^w_{\bar{d}}$ が与えられた際に，学習されたパラメータ $\hat{\eta}$，$\hat{\beta}$，$\hat{\theta}$，を固定し，前述のアルゴリズムにおけるモデルのサンプリングとカテゴリのサンプリングを，未知物体 $\bar{d}$ に対してのみ行う．モデル $l$ における未知物体 $\bar{d}$ のモダリティ $m$ の特徴量 $w^m_{\bar{d}}$ がカテゴリ $k^l$ である確率は，以下のように計算する．

$$P(w^m_{\bar{d}}|\hat{\mathbf{W}}^l_k) = \frac{\hat{N}^{lm}_{k\bar{w}^m} + \alpha^{lm}}{\hat{N}^{lm}_k + d^m\alpha^{lm}} \tag{8}$$

ただし，$\hat{\mathbf{W}}^l_k$ は，学習時にカテゴリ $k^l$ に属する特徴量の集合である．また，未知物体のクラス $t^l$ に属する特徴量 $\bar{\mathbf{W}}^l_t$ に割り当てられるカテゴリ $\bar{k^l_t}$ は次の式からサンプリングする．

$$\bar{k^l_t} \sim P(\bar{\mathbf{W}}^l_t|\hat{\mathbf{W}}^l_k) \tag{9}$$

すなわち，各モデルにおいて，カテゴリ数を固定し，クラスの選択のみを行うことになる．最終的にモデル $l$ における未知物体 $\bar{d}$ のカテゴリは，学習時と同様に式 (7) より計算可能である．さらに，モデルのサンプリングに関しては，式 (1) の代わりに以下の式を用いる．

$$P(l_{\bar{dn}}|\hat{\mathbf{W}}^{lm}, w^m_{\bar{dn}}, \delta) \propto P(l_{\bar{dn}}|\delta)P(w^m_{\bar{dn}}|\hat{\mathbf{W}}^{lm}) \tag{10}$$

パラメータを固定し，モデルのサンプリング及びカテゴリのサンプリングを繰り返し行うことで，各モデルにおける未知物体 $\bar{d}$ のカテゴリの推定が可能となる．

### 2.5 未観測情報の予測

IMoMs の有用性は，観測された一部の情報から，他の未観測情報を予測できる点にある．すなわち，ロボットは物体を見るだけで，それがどのような硬さなのか，どのような音を発するのかを予測できる．さらに，単語情報を予測することで，

図 3: 実験に使用した 67 個の物体

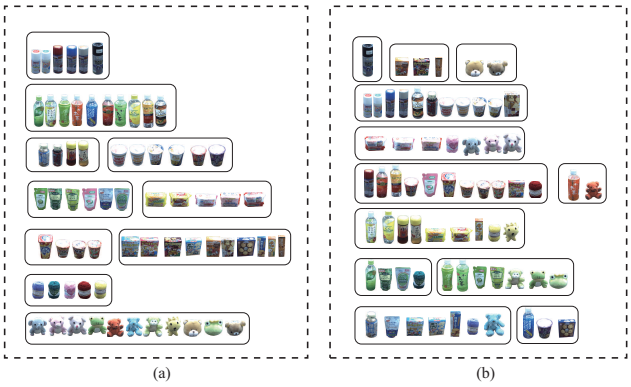


図 4: IMoMHDPs によるカテゴリ分類結果

得られた情報を単語で表現することも可能となる．逆に，単語が与えられると，そこからマルチモーダル情報を予測することで，単語の意味を理解することができる．ここでは，未知物体 $\bar{d}$ の視覚情報 $w^v_{\bar{d}}$ から，単語情報 $w^w_{\bar{d}}$ を予測することを考える．まず，前節と同様にモデルのサンプリングと，学習時に形成された各モデル $l$ における物体のカテゴリ推定を行い，最終的に，以下の式より単語 $w^w_{\bar{d}}$ が発生する確率を求めることができる．

$$\begin{aligned} P(w^w_{\bar{d}}|w^v_{\bar{d}}) &= \sum_l P(w^w_{\bar{d}}|w^v_{\bar{d}}, l)P(l|\delta) \\ &= \sum_l \sum_k P(k|w^v_{\bar{d}}, l)P(w^w_{\bar{d}}|\hat{\mathbf{W}}_k, l)P(l|\delta) \end{aligned} \tag{11}$$

ここでは単語の予測のみについて述べたが，他の予測も同様に可能である．

## 3. 実験

提案手法の有効性を示すため，図 2 に示すロボットにより，図 3 に示す 67 個の物体から取得した視覚・聴覚・触覚・単語情報を用いて実験を行った．

### 3.1 IMoMs によるカテゴリ分類

まず，図 3 に示す 67 個の物体に対して，IMoMHDPs によりカテゴリ分類を行った．その結果を図 4 に示す．図 4 に示されているように，IMoMHDPs により推定された概念構造の数は 2 つとなり，図 4(a) の概念ではスプレー缶，ペットボトル，ビン，毛糸など，ガラガラとぬいぐるみが 1 つのカテゴリに分類されているものの，物体を表す物体概念が正しく形成できている．また，図 4(b) の概念では黒，茶色，白，ピンク色，赤，オレンジ色，黄色など，青いスプレー缶やねずみ色の象のぬいぐるみが誤った分類が行われているが，色を表す概念が正しく形成できている．視覚 (SIFT)・触覚・聴覚情報など，概念形成に用いているマルチモーダル情報の多くは物体概念を形成することに寄与すると考えられるが，物体概念では表現できない色の情報は異なるモデルに分類され，色概念が形成された考えられる．

さらに，各概念に割り振られていた各モダリティの割合を，図 5 に示す．図 4(a) の概念は物体概念を表現しているため，視覚 (SIFT)・聴覚・触覚情報の割合が高く，色情報の割合が低いことが分かる．また，ロボットに教示された単語の多くは物体名などの名詞であるため，単語情報の割合も多く占めてい

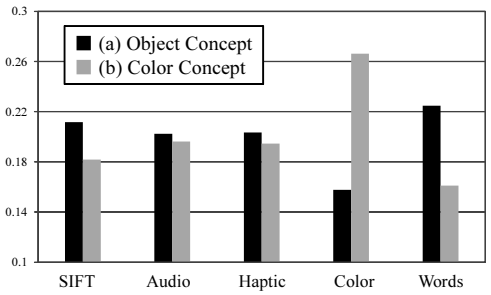


図 5: 各概念に割り振られた情報の割合

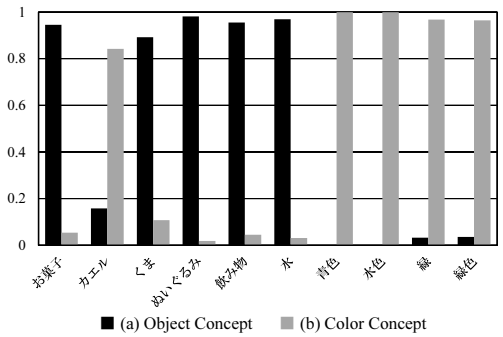


図 6: 各概念に割り振られた単語の割合

る．一方，図 4(b) の概念は色概念を表現しているため，他のモダリティの情報に比べて色情報が多くの割合を占めていることがわかる．

さらに，図 6 に，各概念に割り振られた単語情報の一部を示す．図 6 から分かるように，色を表現する単語が正しく色情報を多く用いて形成された概念である図 4(b) に割り振られていることが分かる．しかしながら，「カエル」は物体名を表現する単語ではあるが，色概念（図 4(b)）に割り振られている．これは学習に用いた物体の中に，“緑色のカエル” しか存在しておらず，“カエル” は “緑色” と同じく色を表す単語として学習されたためである．実際，様々な色が存在する “くま” は物体概念（図 4(a)）に割り振られている．また，“青”，“青色” といった単語はどちらも色を表現するため，色概念（図 4(b)）に割り振られているのに対し，“水”，“水色” といった単語では，“水” が “飲み物” と同じ物体を表す単語として学習されたのに対して，“水色” は色概念を表す単語として学習されている．以上より，単語情報が正しく概念に接地できていると言える．

### 3.2 単語情報の予測

次に，認識用物体の視覚・聴覚・触覚情報から，単語情報の予測を行った．なお，比較のため従来手法である MHDP でも同様の予測を行った．実験の結果，IMoMHDPs の単語予測精度が 90.7% に対して，MHDP は 77.8% となり，単語の予測精度が 13%ほど向上した．なお，各物体から予測された単語数の平均は IMoMHDPs が 3.9 単語，MHDP は 4.1 単語であり，大きな差はなかった．

さらに，図 7 に実際に予測された単語の一例を示す．図 7 のペットボトルにおいて，MHDP では “柔らかい”，“毛糸” といった単語が高い確率で想起されている．これは，学習に用いている物体に「青色のペットボトル」が含まれておらず，全てのモダリティの情報を統合して 1 つの概念を形成する MHDP では，様々な色を持つペットボトルを表現するカテゴリが上手く形成できいなかったためだと考えられる．その結果，未知の青いペットボトルを誤って “毛糸” と認識したと考えられる．IMoMs においても，ペットボトルのことを “お菓子” と誤った単語を予測しているが，これは硬さ，振った際の音などの情

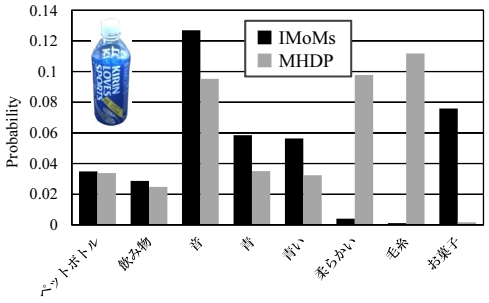


図 7: 未知物体から予測された単語例

報が似通っているためだと考えられる．

## 4. まとめ

本稿では，ロボットが取得した視覚・聴覚・触覚・単語情報を分類することで様々な概念を形成可能な手法 IMoMs を提案した．提案手法によって，注目する特徴が異なる複数の概念構造を教師なしで形成することが可能であり，ロボットは物体概念だけでなく，物体概念では表現のできない色を表現する概念など，複数の概念を同時に獲得することが可能となった．限定された状況ではあるものの，実験を通して実際に複数の概念が獲得されていることを示した．さらに興味深いことは，概念と単語との結びつきである．MHDP により形成される概念構造は 1 つであるため，物体名である「ぬいぐるみ」や形容詞の「青い」といった単語が一つの概念に結びつくのに対して，IMoMHDPs では，「ぬいぐるみ」と「青い」がそれぞれ別の概念に結びつくこととなり，より人間の感覚に近い単語と概念の関係性が学習できていている．IMoMs のもう一つの重要な利点は，ノンパラメトリックベイズ手法を用いているために，データの複雑さにあわせて自動的に概念の数を推定できることにある．

今後の課題としては，学習のオンライン化が挙げられる．また，視覚・聴覚・触覚・単語情報だけではなく，物体を扱う動作など，分類に用いるモダリティを増加させることが，さらなる多種多様な概念形成に必要だと考えている．

## 参考文献

[Aldous 85] Aldous, D.: Exchangeability and related topics, *École d'Été de Probabilités de Saint-Flour XIII-1983*, pp. 1–198 (1985)

[Araki 12] Araki, T., Nakamura, T., Nagai, T., Funakoshi, K., Nakano, M., and Iwahashi, N.: Online object categorization using multimodal information autonomously acquired by a mobile robot, *Advanced Robotics*, Vol. 26, No. 17, pp. 1995–2020 (2012)

[Lowe 04] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110 (2004)

[Nakamura 12] Nakamura, T., Araki, T., Nagai, T., and Iwahashi, N.: Grounding of Word Meanings in LDA-Based Multimodal Concepts, *Advanced Robotics*, Vol. 25, pp. 2189–2206 (2012)

[中村 10] 中村 友昭, 西田 匡志, 長井 隆行 : 把持動作による物体カテゴリの形成と認識, 情報処理学会全国大会, 5V-3 (2010)

[中村 13] 中村 友昭, 荒木 孝弥, 長井 隆行, 岩橋直人 : 階層ディリクレ過程に基づくロボットによる物体のマルチモーダルカテゴリゼーション, 計測自動制御学会論文集, Vol. 49, No. 4, pp. 469–478 (2013)