

多層マルチモーダルLDAを用いた多様な概念の統合と語意の獲得

Integration of Various Concepts and Words Acquisition Using Multi-layered Multimodal LDA

ファドリル ムハンマド アッタミミ ムハンマド 阿部香澄 中村友昭 長井隆行
 Muhammad Fadlil Muhammad Attamimi Kasumi Abe Tomoaki Nakamura Takayuki Nagai

電気通信大学情報理工学研究所

Graduate School of Informatics and Engineering, The University of Electro-Communication

In the field of intelligent robotics, object handling by robots can be achieved by capturing not only the object concept through object categorization, but also other concepts (e.g., the movement while using the object), as well as the relationship between concepts. Moreover, capturing the concepts of places and people is also necessary to enable the robot to gain real-world understanding. In this study, we propose multi-layered multimodal latent Dirichlet allocation (mMLDA) to realize the formation and integration of various concepts by robots. Another issue to be addressed is the language acquisition by the robots. We propose a method to infer which words are originally connected to a concept using mutual information between words and concepts. Moreover, the order of concepts in teaching utterances can be learned using a simple Markov model, which corresponds to grammar. This grammar can be used to generate sentences that represent the observed information.

1. はじめに

知能ロボットの要素技術として物体のカテゴリ分類・認識があり、未知の環境で柔軟に動作するためにもこうした技術が重要である。著者らのグループでは、Latent Dirichlet Allocation (LDA)をはじめとする統計モデルを自律型ロボットに応用することで、物体の概念形成と語意獲得を行う手法を提案してきた。これらは、ロボットが経験することによって取得するマルチモーダルな知覚情報のカテゴリ分類に基づいており、LDAをマルチモーダル情報に拡張したマルチモーダルLDA (MLDA) [Nakamura 07] や、マルチモーダル hierarchical Dirichlet process (MHDP) によって実現される。また、我々が用いている言語もカテゴリに基づいており、ロボットがカテゴリ分類を通じて物体の概念を獲得することで、未観測情報の予測や言語の理解が可能になると考えられる。しかし、ロボットが知的にふるまうためには物体のカテゴリ認識や予測だけでは不十分であり、物体と動作など異なる種類の概念間の関係性を獲得する必要がある。

本稿では、物体、動き、場所、人物といった複数の概念とその関係性を表現するためのモデルである多層MLDA (mMLDA) を提案する。物体概念は、物体をロボットが観測し、得られるマルチモーダル情報をMLDAで分類することで形成する。また、動き概念はロボットに搭載されたKinectから、物体を扱っている人の関節角を取得し、これらの情報をLDAで分類することで形成する。場所概念や人の概念も同様に、ロボットが取得する座標や、顔画像処理技術を用いて推定する年齢・性別の情報をMLDAで分類することで形成する。さらにmMLDAでは、これら複数の概念を統合するMLDAを上層に配置し、これらの関係を表す上位概念を形成することになる。すなわち提案モデルは、多層のMLDAで構成されており、下層のMLDAでは物体、動き、場所、人物の概念(下位概念)がそれぞれ形成され、上層のMLDAでこれらの概念を統合している(上位概念)。ここでのポイントは、これらの下位概念と上位概念が同時に教師なし学習されることである。このモデルでは、例えば下位概念としてジュースという物体概念や物を口に運ぶ動きの概念、ダイニングという場所の概念などが形成される。上位層ではこれらの関係性が学習され、「飲む」という行動概念が形成される。これにより、ジュースを見ることでそれを口に運ぶ「飲む」という行動や、その「飲む」という行動が「ダ

イニング」という場所で行なわれやすいといった未観測情報の予測を行うことが可能となる。

また、言語理解・生成のための語意獲得についても、mMLDAを用いて実現することが可能である。これは、[Nakamura 09]で提案されているように、単語をBag of Words (BoW) 表現とし、モダリティーの一つと考えることで解決することができる。しかし階層のないMLDAと異なる点は、どの階層のどの概念クラスにどの単語が結びつくかを考える必要がある点である。この情報は教示発話には明示的に含まれていないため、学習アルゴリズムが何らかの基準に従ってこの結びつきを見出す必要がある。本稿では、単語と概念間の相互情報量を用いることで、どの単語が本来どの概念クラスに結びついているのかを自動的に推定する手法を提案する。さらに、単語と概念の結びつきを用いて、教示発話における概念クラスの生起順を学習することが可能であり、これを文法と捉えることで、観測した情報を表現する文章を生成することができる。

関連研究として、物体のカテゴリ分類に関する研究や動作のモデル化に関する研究 [濱畑 10] を挙げることができる。しかしこれらの研究では、物体の分類や動作のモデル化など単一の概念のみを考えており、それらの関係性は考えられていない点で本研究とは異なる。また [野田 13] では、Deep Learningを用いた感覚運動情報を統合化する学習フレームワークを提案している。その結果、ロボットの行動パターンと物体や物体の運動が統合的に学習され、例えばロボットの運動から視覚情報を予測することが可能であることを示している。しかし、場所や人といった概念を扱っていない点や、言語情報などを扱っていない点、物体のカテゴリなどを陽には扱っていない点において本研究とは異なっている。また、物体数や運動のパターンが少なく、どれほどの一般性を有しているかが必ずしも明らかではない。

2. 提案手法

ここでは、MLDAを用いて形成された物体、動き、場所、人物の概念を統合することで、より上位の概念を形成することを考える。図1に提案するmMLDAのグラフィカルモデルを示す。図1において、 z は統合概念を表すカテゴリであり、 z^O , z^M , z^P , z^U はそれぞれ下位概念に相当する、物体、動き、場所、人物カテゴリである。上位カテゴリ z は、下位カテゴリ間の関係性を表現したモデルとなっている。また、 w^v , w^a , w^h は、それぞれ物体から得られる視覚、聴覚、触覚情報であり、 w^p , w^c , w^s , w^y は物体を扱っている際の人の動き、座標、性別、

連絡先: 長井隆行, 電気通信大学大学院情報理工学研究所,
 東京都調布市調布ヶ丘 1-5-1, tngai@ee.uec.ac.jp

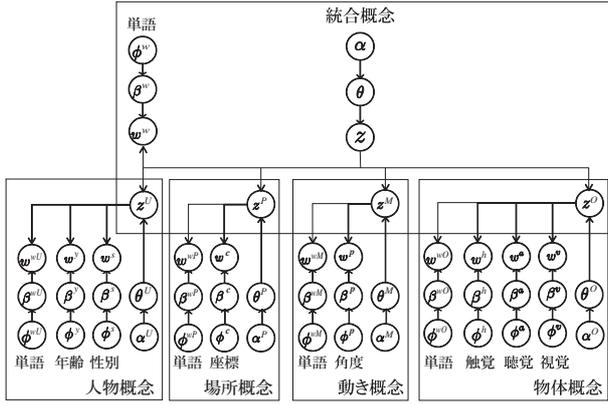


図 1: 多層マルチモーダル LDA のグラフィカルモデル

年齢の情報である。さらに、 $w^w, w^{wO}, w^{wM}, w^{wP}, w^{wU}$ は、教示発話から得られる単語情報である。以下それぞれの概念について詳しく述べる。

2.1 下位概念

2.1.1 物体概念

物体概念は、ロボットが実際に取得したマルチモーダル情報をカテゴリ分類することにより形成する。つまり、視覚・聴覚・触覚・単語情報 w^v, w^a, w^h, w^{wO} がその類似性により分類され、物体カテゴリ z^O が形成される。知覚情報は Bag of Features (BoF) モデルを用いて表現する。視覚情報としては、取得した物体画像を 128 次元の DSIFT を用いて変換し、ベクトル量子化することで 500 次元のヒストグラムとする。聴覚情報は、MFCC を特徴量とし、ベクトル量子化することで 50 次元ヒストグラムとする。同様に、触覚情報は、取得した触覚センサーのデータに対する曲線近似のパラメータをベクトル量子化し、15 次元ヒストグラムとする。また、全ての教示発話を単語分割し、BoW モデルを用いて表現したものを単語情報として扱う。

2.1.2 動き概念

前述の物体概念と同様に、図 1 の下側に示す MLDA と等価なモデルによって動き概念の形成を行なう。動き情報は、人の動作中の 11 箇所関節角度を、動作開始から動作終了まで Kinect を用いて取得することを前提とする。また動きの情報は、操作対象となる物体によって分節できると仮定している。1 つの動作から複数の 11 次元の特徴ベクトルが得られ、それをあらかじめ計算した 70 の代表ベクトルによりベクトル量子化することで 70 次元のヒストグラムとし、これを動き情報として用いる。

2.1.3 場所概念

場所概念は、人の動作中の座標を動作開始から動作終了まで取得することで形成する。1 つの動作から複数の 2 次元座標が得られるため、これらをベクトル量子化し、6 次元のヒストグラムとすることで場所情報とする。代表ベクトルは、学習データを k 平均法によりクラスタリングすることで決定する。

2.1.4 人物概念

人物概念の形成では、動作中の人の顔画像から、性別及び年齢の推定を行い、これらの値を人物情報として扱う。他の概念と同様、性別・年齢の推定結果を基にデータの量子化を行い、2 次元の性別ヒストグラムと 10 次元の年齢ヒストグラムを人物情報として用いる。

2.2 統合概念

提案モデルにおいて、物体、動き、場所、人物概念は、それぞれ MLDA で表現されており、さらにその上位でそれらの概念の関係を統合概念としての MLDA で表現する二層構造と

なっている。物体、動き、場所、人物概念を独立した MLDA として学習し、その学習結果を統合概念として学習することも可能であるが、そのような方法では各概念が相互に影響を与えることができない。そこで、各概念を表す隠れ変数 $z, z^C \in \{z^O, z^M, z^P, z^U\}$ を同時に学習する。学習にはギブスサンプリングを用い、各概念を表すカテゴリ z, z^C を、観測データ $w^m \in \{w^v, w^a, w^h, w^{wO}, w^p, w^{wM}, w^c, w^{wP}, w^s, w^y, w^{wU}, w^w\}$ からサンプリングすることで学習する。サンプリングは、 $\theta, \theta^C, \beta^m$ を周辺化した以下の事後分布を用いる。

$$P(z_{jmi}, z_{jmi}^C | \mathbf{Z}_{-jmi}, \mathbf{Z}_{-jmi}^C, \mathbf{W}^m) \propto P(z_{jmi} | \mathbf{Z}_{-jmi}) P(z_{jmi}^C | z_{jmi}, \mathbf{Z}_{-jmi}, \mathbf{Z}_{-jmi}^C) \times P(w_{ji}^m | z_{jmi}^C, \mathbf{Z}_{-jmi}^C, \mathbf{W}_{-ji}^m) \quad (1)$$

右辺のそれぞれの確率分布は次のようになる。

$$P(z_{jmi} = k | \mathbf{Z}_{-jmi}) = \frac{\alpha + N_{j,z=k}^{-jmi}}{K\alpha + N_j^{-jmi}} \quad (2)$$

$$P(z_{jmi}^C = l | z_{jmi} = k, \mathbf{Z}_{-jmi}, \mathbf{Z}_{-jmi}^C) = \frac{\alpha^C + N_{z=k, z^C=l}^{-jmi}}{K^C \alpha^C + N_{z=k}^{-jmi}} \quad (3)$$

$$P(w_{ji}^m = x | z_{jmi}^C = k, \mathbf{Z}_{-jmi}^C, \mathbf{W}_{-ji}^m) = \frac{\phi^m + N_{z^C=k, w^m=x, m}^{-jmi}}{W^m \phi^m + N_{z^C=k, m}^{-jmi}} \quad (4)$$

ただし、 \mathbf{Z}, \mathbf{Z}^C は、それぞれ全物体の全情報に割り当てられた上位カテゴリ、下位カテゴリの集合を表し、 \mathbf{W}^m はモダリティ m の全物体の情報の集合である。 N_{jz} は物体 j の全モダリティに上位カテゴリ z が割り当てられた回数であり、 N_{z^C, w^m} はモダリティ m の特徴量 w^m に下位カテゴリ z^C が割り当てられた回数である。また、 N_{z, z^C} は上位カテゴリ z と下位カテゴリ z^C の共起した回数を表しており、 K, K^C, W^m はそれぞれ上位カテゴリのカテゴリ数、概念 C のカテゴリ数、モダリティ m の情報の次元数である。負の添字はその情報を除外することを表し、 $-jmi$ は j 番目の物体のモダリティ m の i 番目の情報を除外することを表している。

モデルの学習は、隠れ変数である z, z^C を、収束するまで事後分布からサンプリングすることによって実現できる。以上のようなサンプリングを繰り返すことで、 N_* がある値へと収束する。 K を上位カテゴリのカテゴリ数とする時、最終的なパラメータの推定値 $\hat{\beta}_{w^m z^C}^m, \hat{\theta}_{z z^C}^C, \hat{\theta}_{jz}$ は以下ようになる。

$$\hat{\beta}_{w^m z^C}^m = \frac{N_{z^C, w^m} + \phi^m}{N_{z^C, m} + W^m \phi^m} \quad (5)$$

$$\hat{\theta}_{z z^C}^C = \frac{N_{z z^C} + \alpha^C}{N_z + K^C \alpha^C} \quad (6)$$

$$\hat{\theta}_{jz} = \frac{N_{jz} + \alpha}{N_j + K\alpha} \quad (7)$$

さらに、学習モデルを用いることで、物体や動きの認識だけでなく、概念間の予測も可能となる。

2.3 近似モデル

上述した提案モデル以外にも、各概念を MLDA により独立に形成し、フィードフォワードに接続することで、簡易的に物体、動き、場所、人物、統合概念を形成できる。つまり、物体、動き、場所、人物概念を学習した後、統合概念 z のみを学習することになる。しかし、後に示す実験の結果から分かるように、各概念を独立的に学習することで、下位概念での学習誤りがそのまま上位概念の学習に影響を及ぼし、モデル全体の精度を下げてしまうことになる。

表 1: 動き, 物体, 場所, 人物データの対応表 (カッコ内の数字はカテゴリ ID)

動き	物体	場所	人物	動き	物体	場所	人物					
持ち上げる (1)	茶碗 (13)	ダイニング (4)	全員 (1,2,3,4)	抱く (8)	ぬいぐるみ (2)	リビング (2)	子供の女 (3)					
	飲み物 (缶) (17)			積み重ねる (9)	積み木 (32)		子供 (3,4)					
	カップヌードル (21)			置く (10)	消臭剤 (7)		大人の女 (1)					
	プラスチックカップ (25)				除湿剤 (8)		子供 (3,4)					
上に投げる (2)	スプレー缶 (23)	庭 (5)	男性 (2,4)	積み木 (32)	ダイニング (4)	リビング (2)	女性 (1,3)					
	ぬいぐるみ (2)	リビング (2)	子供 (3,4)	プラスチックカップ (25)								
	マラカス (29)			手に塗る (11)				ハンドクリーム (6)				
口に運ぶ (3)	ボール (31)	ダイニング (4)	全員 (1,2,3,4)	取り出す (12)	ティッシュ箱 (10)	ダイニング (4)	リビング (2)	全員 (1,2,3,4)				
	金属の食器 (12)			クッキー (20)	大人の女 (1)							
	飲み物 (缶) (17)			フロリングワイパー (3)	大人の女 (1)							
	ペットボトル (18)			野菜 (玩具) (27)	キッチン (3)			ダイニング (4)	ダイニング (4)	全員 (1,2,3,4)		
	プラスチックカップ (25)			ドレッシング (14)	蜂蜜 (15)							
	茶碗 (13)			ソース (16)	中身を注ぐ (15)			シャンプー (9)	浴室 (6)	大人 (1,2)		
	野菜 (玩具) (27)			スナック (19)	リビング (2)			じょうろ (24)	庭 (5)	大人の男 (2)		
	カップヌードル (21)			車 (玩具) (28)				飲み物 (缶) (17)	ダイニング (4)	全員 (1,2,3,4)		
	右に動かす (4)			フロリングワイパー (3)	ダイニング (4)			大人の女 (1)	ペットボトル (18)	ダイニング (4)	リビング (2)	大人女 (1)
				スポンジ (4)	キッチン (3)			ラップ (11)				
皿を洗う (5)	たわし (5)	ダイニング (4)	大人の女 (1)	塗る (17)	スプレー缶 (23)	庭 (5)	大人の男 (2)					
	ガラガラ (1)			リビング (2)	子供 (3,4)	履く (18)	靴 (30)	玄関 (1)	全員 (1,2,3,4)			
上下に振る (6)	マラカス (29)	ダイニング (4)	全員 (1,2,3,4)	袋を開ける (19)	スナック (19)	リビング (2)						
	ドレッシング (14)											
	ソース (16)											
	飲み物 (缶) (17)											
	ペットボトル (18)											
	スプレー缶 (23)			庭 (5)	大人の男 (2)							
	すくう (7)			ショベル (26)								

2.4 相互情報量を用いた単語の予測

本稿では、図 1 に示したように、各概念に教示発話から得られる全ての単語情報を与えて学習を行う。各概念を表現する適切な単語が存在すると考え、ここで、単語とカテゴリの結び付きの強さの尺度として、単語とカテゴリ間の相互情報量を用いる。単語 w^w と概念クラス $i \in \{ \text{物体概念, 動き概念, 場所概念, 人物概念, 統合概念} \}$ のカテゴリ k との間の相互情報量は以下の式となる。

$$I(w^w, k|i) = \sum_{K,W} P(W, K|i) \log \frac{P(W, K|i)}{P(W|i)P(K|i)} \quad (8)$$

但し、 $K \in (k, \bar{k})$, $W \in (w^w, w^{\bar{w}})$ とし、 \bar{k} は k 以外のカテゴリを表す。また、 $w^{\bar{w}}$ は w^w 以外の単語を表している。相互情報量とは、二つの確率変数の共有する情報量であり、相互依存の尺度である。したがって単語とカテゴリ間の相互情報量が大きい場合、その単語はそのカテゴリを表現しているといえる。

本稿では、単語によって、複数の概念を表す可能性があると考え、式 (8) を用いて求めた相互情報量を単語の各概念クラスに対する重みとして考える。その重みを $weight(i, w^w)$ とし、次式で単語予測スコアを計算する。

$$weight(i, w^w) = \max_k I(w^w, k|i) \quad (9)$$

$$\hat{P}_i(w^w | w_{obs}^m) = weight(i, w^w) P_i(w^w | w_{obs}^m) \quad (10)$$

このように、単語の各概念クラスに対する重みを求め、概念クラス i の単語予測 $P_i(w^w | w_{obs}^m)$ の際に重みをつけることで、各概念から生成される単語の予測精度を向上させることが可能となる。

3. 文章生成

3.1 文法の学習

mMLDA を用いることで、観測情報を表現するのに適切な単語を予測することができる。文章を生成するためには、さらに文法を考える必要がある。本稿では、mMLDA における概念クラス (式 (10) における i) の発火順を文法と考える。これは、各単語に対する概念選択により、単語は特定の概念クラスと結びつけることができるためであり、教示発話を単語分割し

その各単語の概念クラスを推定することで実現できる。但し、ここでは助詞や機能語を考えないこととする。

例えば、“母はキッチンで野菜を切る” という発話から概念クラスを、“母—人物; キッチン—場所; 野菜—物体; 切る—動き;” と推定することができ、結果的に“(人物) (場所) (物体) (動き)”なる文法が得られる。ここでは文法を、マルコフモデルで表現することとし、学習データである教示発話から次のように学習する。

$$P(C_t | C_{t-1}) = \frac{N_{C_{t-1}C_t}}{N_S} \quad (11)$$

但し、 C_t は文章中の t 番目の単語に該当する概念クラスである。また、 $N_{C_{t-1}C_t}$ と N_S はそれぞれ C_{t-1} から C_t に遷移した回数と概念クラス間遷移の総数である。

3.2 観測情報からの文章生成

ロボットが観測情報を文章で表現するためには、予測した単語を上記の文法を基に並べればよい。具体的には、文頭から順に $P(C_t | C_{t-1})$ から t 番目の概念クラスである C_t をサンプリングする。そして、 $w_t = \text{argmax}_{w^w} P(w^w | w_{obs}^m, C_t)$ に従って t 番目の単語を計算する。この手順を、 C_t が文末になるまで繰り返す。

4. 実験

提案するモデルの有効性を示すため、表 1 に示すデータを用いて実験を行った。

4.1 下位概念の形成

提案モデルと近似モデルによって形成された物体概念を評価した。分類精度を計算した結果、mMLDA では 74.24%、近似モデルでは 65.15% となり、提案モデルである mMLDA の方がより正解に近い分類ができている。mMLDA では、“飲み物 (缶) (17)” は一つのカテゴリに分類できたのに対して、近似モデルでは、この物体を 3 つのカテゴリに分割してしまった。同じ“飲み物 (缶) (17)”でも、異なる柄や形を持つため、近似モデルでは異なるカテゴリに分類されてしまったのに対して、mMLDA では“飲み物 (缶) (17)”と関係する動きも考慮して分類を行なうため、正しく一つのカテゴリに分類することができたと考えられる。

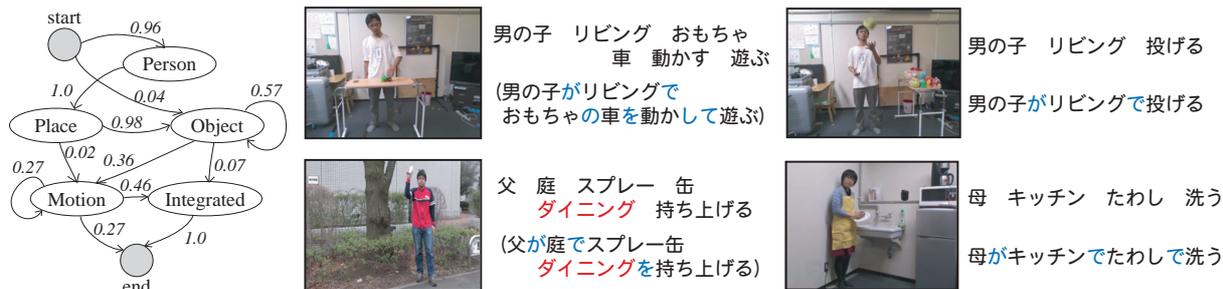


図 3: 獲得した文法 (左) と観測情報から生成された文章の例 (右) 上段: システムの出力, 下段: 人手で助詞を補った文章

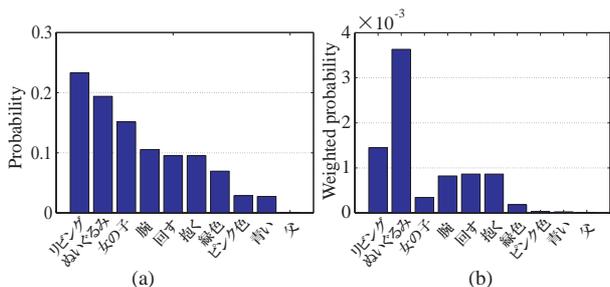


図 2: “ぬいぐるみ”からの単語予測: (a) 単語の発生確率, (b) 相互情報量による重み付けをした単語発生確率

また、動きの種類精度は、mMLDA と近似モデルでそれぞれ、81.06%と75.0%であった。場所概念では、mMLDA と近似モデル共に96.97%であり、人物概念では、それぞれ75.75%と71.21%となった。このように、下位概念の形成結果においては、提案モデルである mMLDA の方がより正解に近い分類ができています。

4.2 統合概念の形成

mMLDA の上位層では物体、動き、場所、人物の関係性を表すカテゴリが形成されており、その中には人にとって意味のあるカテゴリも形成されている。例えば、動きの“口に運ぶ(3)”と物体の“飲み物(缶)(17)”や、“ペットボトル(18)”, “プラスチックカップ(25)”等が1つのカテゴリに分類された。これは、“何かを飲む”という概念を表していると考えられる。さらに、“左右に動かす(4)”が、物体によって異なる上位カテゴリに分類された。1つは“車(玩具)(28)”と関係し、もう一つは“フローリングワイパー(3)”と関係する上位カテゴリであり、これらはそれぞれ“車の玩具を走らせる”という概念と、“フローリングワイパーで掃除をする”といった概念であるといえる。このように同じ動きに対しても、物体によって、意味の異なる統合概念が形成できていると言える。

4.3 未観測情報の予測

未観測情報の予測性能を評価するため、可観測の情報から未観測の概念の予測を行った。表2に mMLDA と近似モデルを用いた未観測予測の結果を示した。mMLDA では、近似モデルに比べ下位概念の関係性を正しく捉えられているため高い精度となっていることが分かる。

次に、物体概念から単語の予測を行った。物体の“ぬいぐるみ”から予測された単語が図2である。図2(a)は“ぬいぐるみ”の視・聴・触覚情報が観測された時の単語の発生確率を表し、場所概念を表す“リビング”という単語が一番高い確率を持つと予測された。一方、本稿で提案した相互情報量を各概念クラスに対する重みづけとして求めた結果が図2(b)である。重みづけの結果から、物体概念を表す“ぬいぐるみ”の単語が一番高いスコアを持つことが分かる。他にも、“スプレー缶”

表 2: 未観測情報の予測精度と単語予測

観測した情報	物体	動き	場所	人物
mMLDA				
視・聴・触覚	-	76.67 %	80.00 %	73.33 %
角度	86.67 %	-	80.00 %	90.00 %
座標	76.67 %	76.67 %	-	100 %
性別・年齢	80.00 %	83.33 %	86.67 %	-
近似モデル				
視・聴・触覚	-	66.67 %	70.00 %	70.00 %
角度	76.67 %	-	73.33 %	80.00 %
座標	70.00 %	76.67 %	-	90.00 %
性別・年齢	76.67 %	73.33 %	80.00 %	-

から予測された単語の発生確率では、“庭”という単語が一番高くなった。この予測結果に単語の相互情報量による重み付けをすることで、“スプレー”と“缶”という単語が正しく予測された。このように、相互情報量の重みづけによって、単語を正しく予測することが可能である。

4.4 文章の生成

提案手法によって獲得した文法を、図3左に示した。また、図右には各シーンにおける観測情報を基に生成した文章を示す。上段がシステムの出力であり、下段が人手で助詞を補った文章である。助詞を考慮していないため不自然ではあるが、シーンを表現しているという意味ではある程度正しい文章が生成されていることが分かる。

5. 結論

本稿では、複数の概念を統合するモデルである mMLDA を提案し、その有効性を示した。これにより、単純な概念を獲得すると共に、それらが関係し合う複雑な概念を獲得できる。さらには、語彙と文法を学習することで、観測情報を簡単な文章で表現できることを示した。今後の課題として、MHDP を階層化することでノンパラメトリックベイズ手法への拡張を行うことが挙げられる。

参考文献

[Nakamura 07] Nakamura, T. et al.: “Multimodal Object Categorization by a Robot”, in Proc. of IROS 2007, pp.2415–2420, 2007.

[Nakamura 09] Nakamura, T. et al.: “Grounding of Word Meanings in Multimodal Concepts Using LDA”, in Proc. of IROS 2009, pp.3943–3948, 2009.

[濱畑 10] 濱畑ほか: “ディリクレ過程と相互情報量による非分節対象物操作のからの動作抽出,” 人工知能学会全国大会, 1J1-OS13-11, 2010.

[野田 13] 野田ほか: “Deep Neural Network を用いたヒューマノイドロボットによる物体操作行動の記憶学習と行動生成”, 人工知能学会全国大会, 2G4-OS-19a-2, 2013.