

言語を獲得するロボットの実現に向けて

電気通信大学 中村 友昭

自己紹介

- ▶ 名前: 中村 友昭
- ▶ 所属: 電気通信大学 助教
- ▶ 研究内容:
 - ▶ 知能ロボティクス, 知能システム
 - ▶ **概念・言語学習ロボット**
 - ▶ 言語理解ロボット
 - ▶ 自律移動ロボット(ロボカップ@ホーム)

The Mobile Manipulation Experiment



×10



研究の目的

- ▶ 人のように言語を獲得するロボットの実現
 - ▶ 人や環境とのインタラクションにより自律的に獲得
 - ▶ 概念形成
 - ▶ 語意の獲得(記号接地)
 - ▶ 語彙の獲得(単語辞書)
 - ▶ 文法の獲得



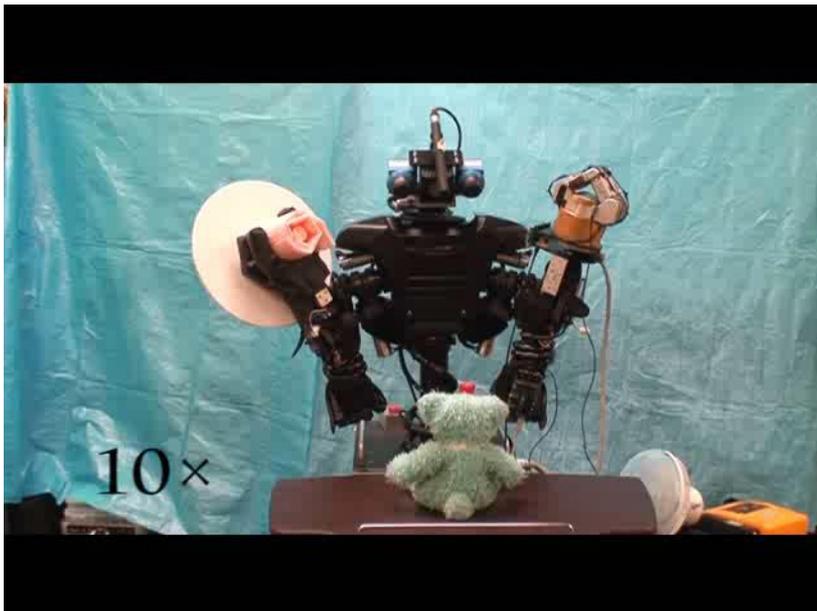
- ▶ 確率モデルを用いて言語獲得アルゴリズムの構築

ロボットによる概念・語意獲得

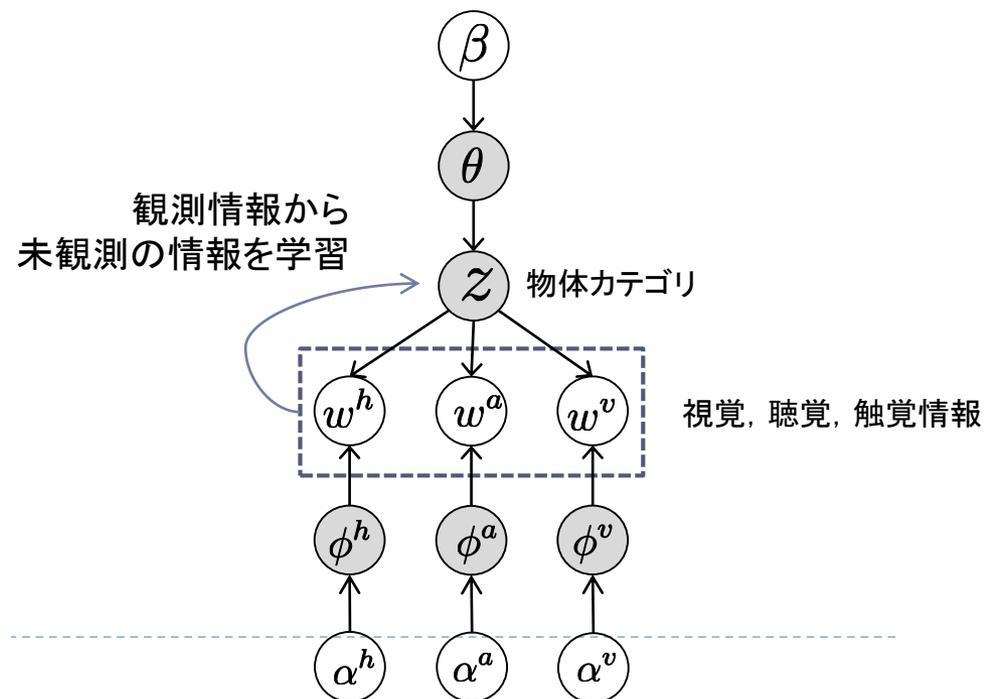
概念形成

▶ 概念の工学的定義

- ▶ 概念 = 知覚情報のクラスタリングによって形成されたカテゴリ
- ▶ 知覚情報: ロボットの視覚, 聴覚, 触覚情報
- ▶ カテゴリ分類: 視覚, 聴覚, 触覚情報の生成モデルを定義



物体に見て・触れることで物体の概念を形成



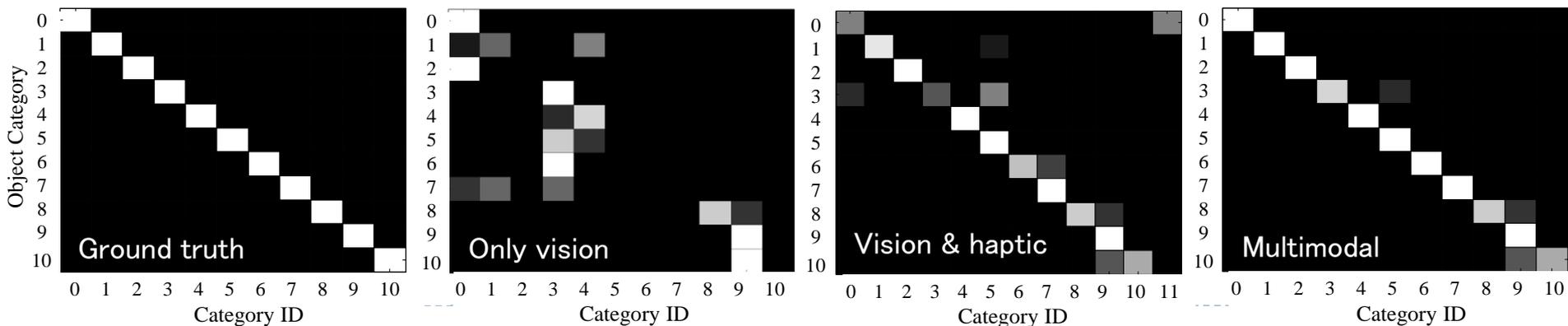
概念形成結果

- ▶ 67個の物体から取得したマルチモーダル情報を分類



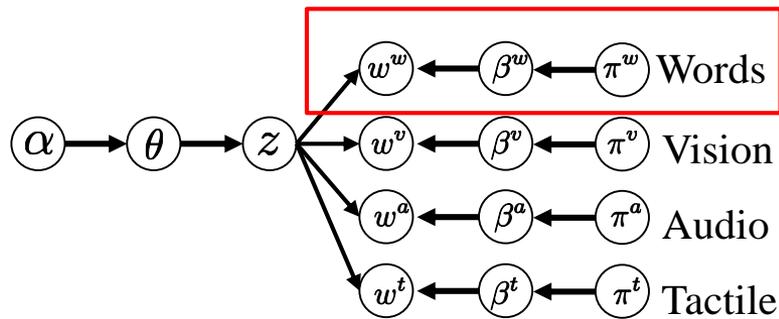
- ▶ 分類結果

(縦軸: 正解カテゴリ, 横軸: 実際に分類されたカテゴリ, 濃淡: 物体の個数)



語意の獲得

- ▶ マルチモーダル情報として単語を追加
- ▶ 単語も含めたマルチモーダル概念を形成



- ▶ 単語からマルチモーダル情報を予測可能



単語を“感覚的”に理解可能
⇒ 単語の意味の理解

ロボットによる語彙の獲得

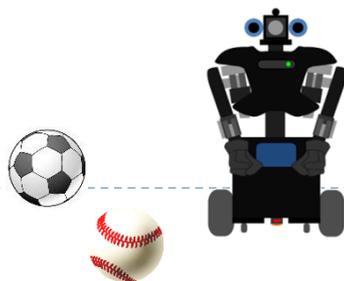
語彙の獲得

- ▶ 前の研究では語彙を持っていることが前提
 - ▶ 語彙＝音声認識・単語分割で使用する単語辞書
- ▶ 語彙を獲得する際の問題
 - ▶ 語彙を持たないため
音声为正しく認識できない
 - ▶ ぼーる？ぼーう？ごーる？
 - ▶ 単語の切れ目が分からない
 - ▶ ぼーる？わぼー？ぼーるだ？るだよ？
- ▶ これらの問題を言語的知識と概念を相互学習することで解決
 - ▶ 言語のパターンに基づく単語分割
 - ▶ 同じ概念に含まれる物体には同じ単語が教示される可能性が高い

これわぼーるだよ



ぼーるがあるよ
これわぼーるだよ



これわ | ぼーる | だよ

言語的知識



相互に学習



概念形成

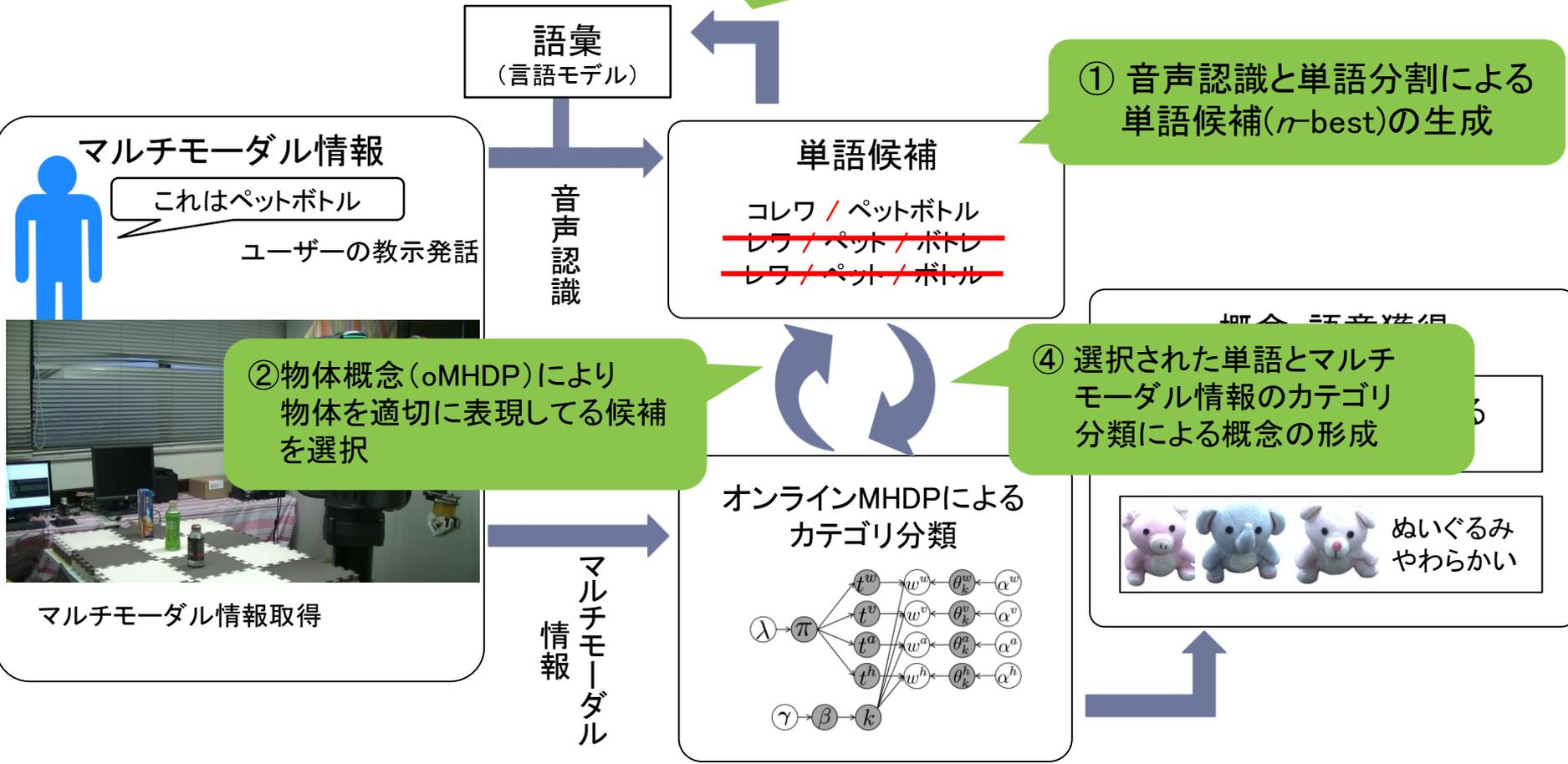
提案手法の概要

③ 選択された認識候補により
語彙(言語モデル)を更新

① 音声認識と単語分割による
単語候補(n -best)の生成

② 物体概念(oMHDP)により
物体を適切に表現してる候補
を選択

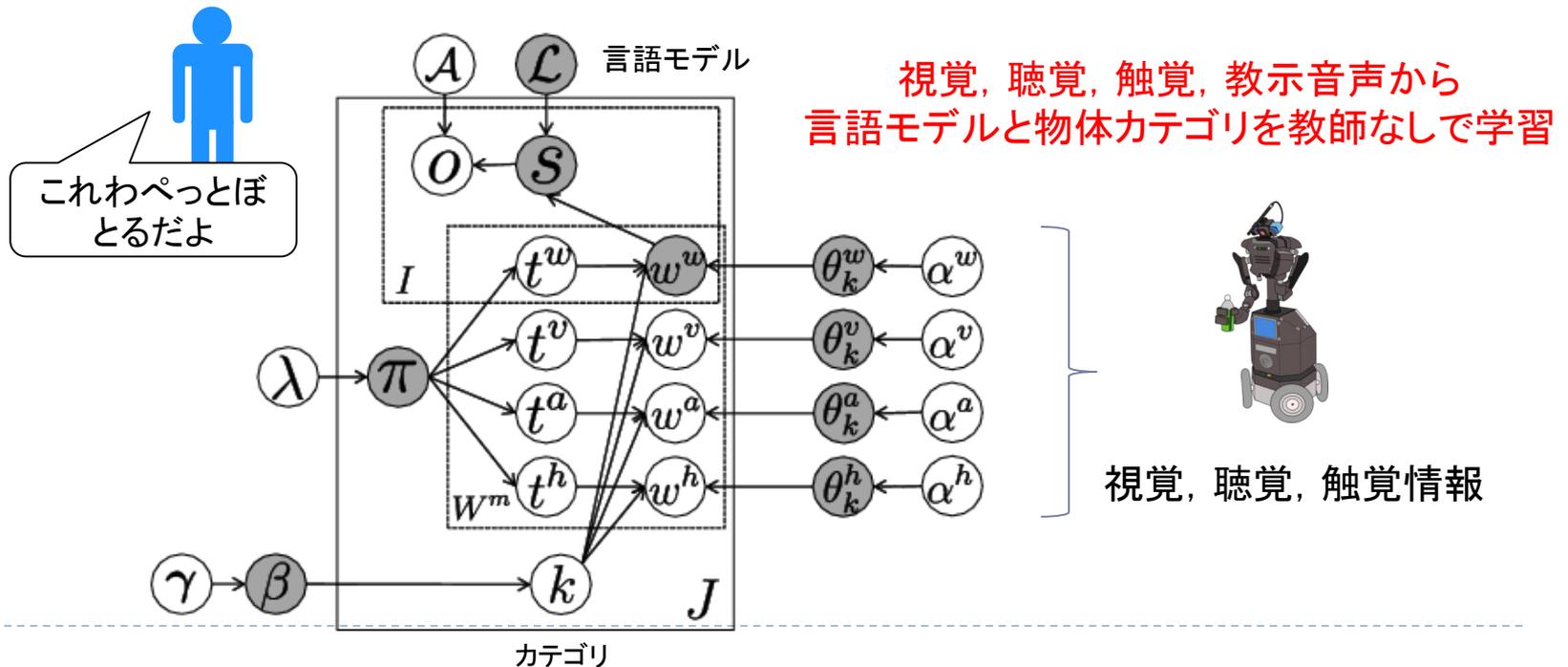
④ 選択された単語とマルチ
モーダル情報のカテゴリ
分類による概念の形成



このプロセスを繰り返すことで概念と言語の相互に学習

提案モデル

- ▶ 教示音声・マルチモーダル情報の生成モデル
 - ▶ 概念の形成と同時に言語モデル(語彙)の獲得が可能
 - ▶ 物体カテゴリ k と言語モデル \mathcal{L} が結びついたモデル → 相互に影響
 - ▶ 同じ物体には同じ単語が発話される可能性が高い
 - ▶ 同じ単語が与えられた物体は同じカテゴリの物体である可能性が高い



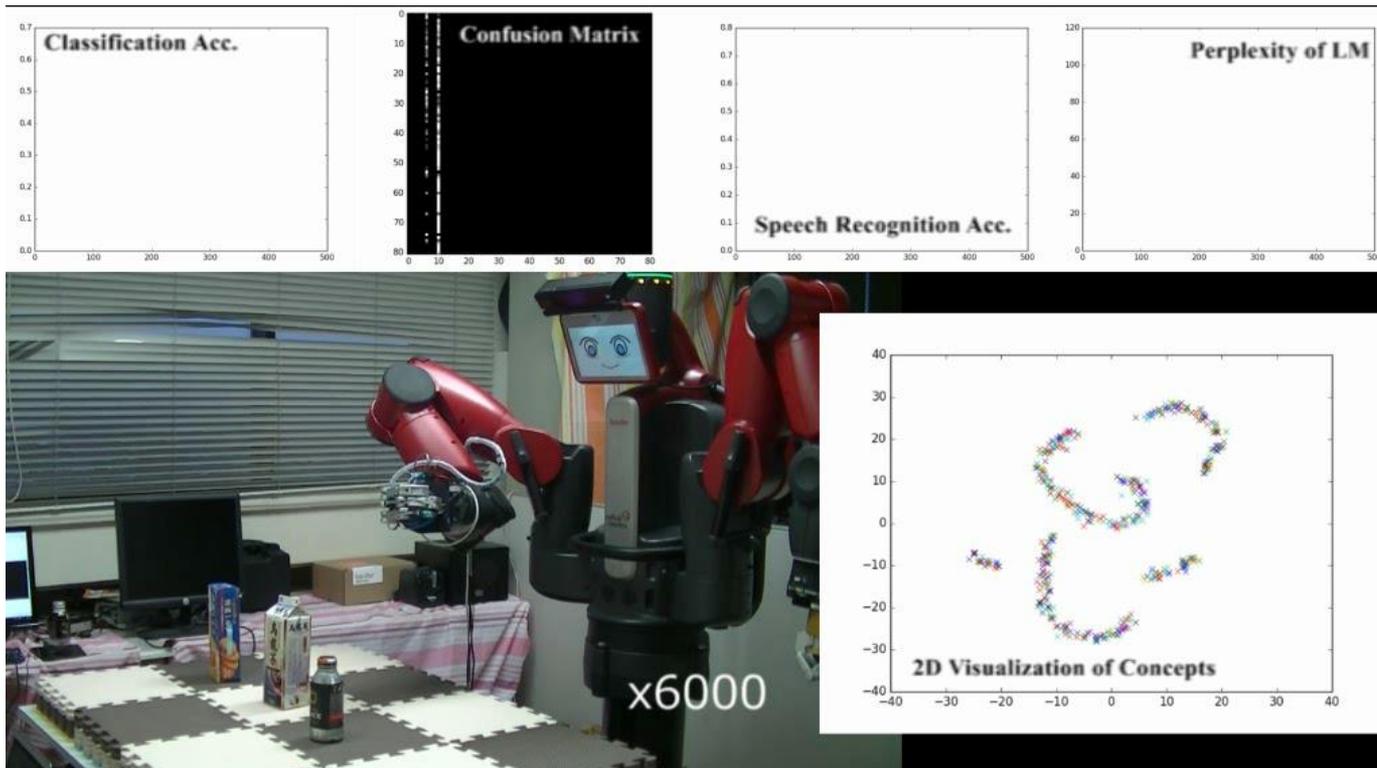
モデルの特徴

- ▶ ノンパラメトリックベイズモデルの導入により
物体の**カテゴリ数を自動的に決定**
- ▶ オンラインで**逐次的に物体を学習**可能
 - ▶ 全ての物体情報を保持することなく1つずつ学習
- ▶ ギブスサンプリングにより, 概念・語彙・単語と概念の結びつきを**教師なしで学習**可能
- ▶ 概念と言語的知識が**相互に影響**し, 双方の精度を向上



学習の様子

- ▶ 1日3～5時間程度の学習を約1ヶ月間実施(100時間強)
- ▶ 計499個の物体を学習



ロボットが獲得した単語

学習初期



約20時間後



約20時間後 (再学習後)



約50時間後



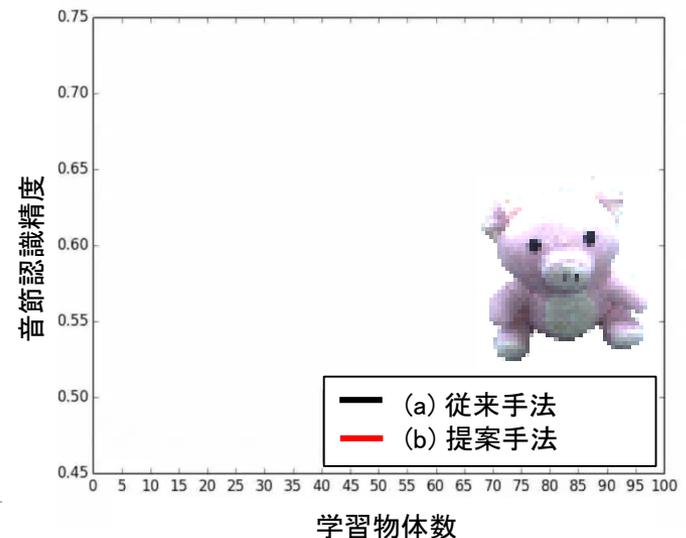
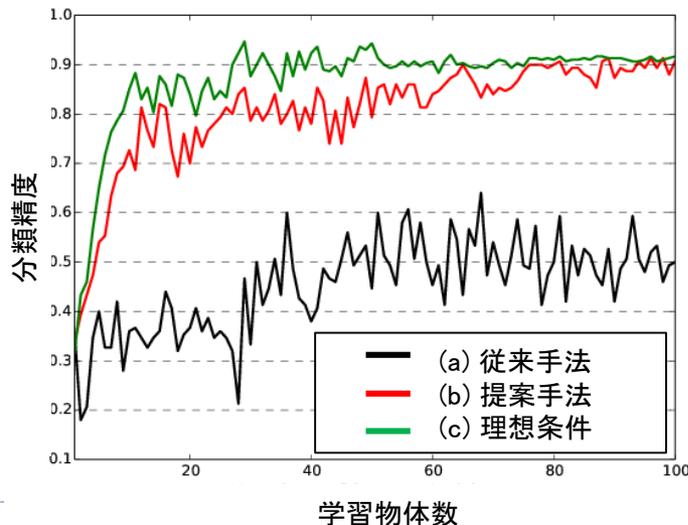
約60時間後



人工知能に発達は必要か？

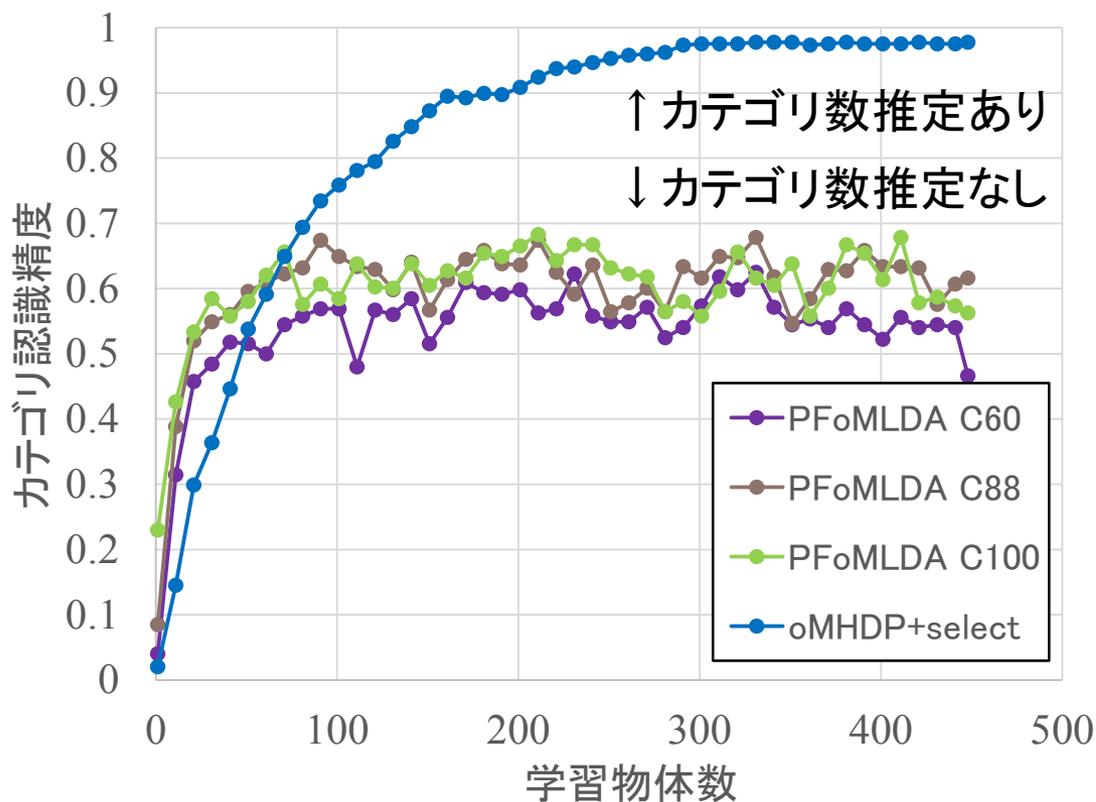
発達が必要か？

- ▶ そもそも発達とは何か？
- ▶ 発達を工学的に定義しないと議論ができない
- ▶ 学習を繰り返すことで徐々に性能が向上すること＝発達
 - ▶ 学習＝個々の事例からモデルのパラメータを更新すること
 - ▶ こう定義すれば本発表での言語獲得ロボットは発達しているといえる？



モデル構造（カテゴリ数）の学習

- ▶ 発達っぽい現象 = 学習によるモデル構造の変化
 - ▶ モデル構造の変化が学習にも影響を与える
- ▶ 言語獲得ロボットはノンパラメトリックベイズにより、オンラインで物体のカテゴリ数を自動的に推定



- 88カテゴリ, 448個の人工データを学習
- カテゴリ数推定のあるモデルとないモデルで分類性能を比較

カテゴリ数推定によって
より精度の高い学習が実現

発達は必要か？

- ▶ あらかじめ正解のカテゴリ数を与えられたモデルですら、学習がうまくいかない
- ▶ 学習初期では得られる情報が少ないにも関わらず、複雑なモデル構造を持っていることが原因
 - ▶ 単純なモデルから複雑なモデルへと徐々に変化（発達？）することで、局所解に陥ってしまうことを防いでいる

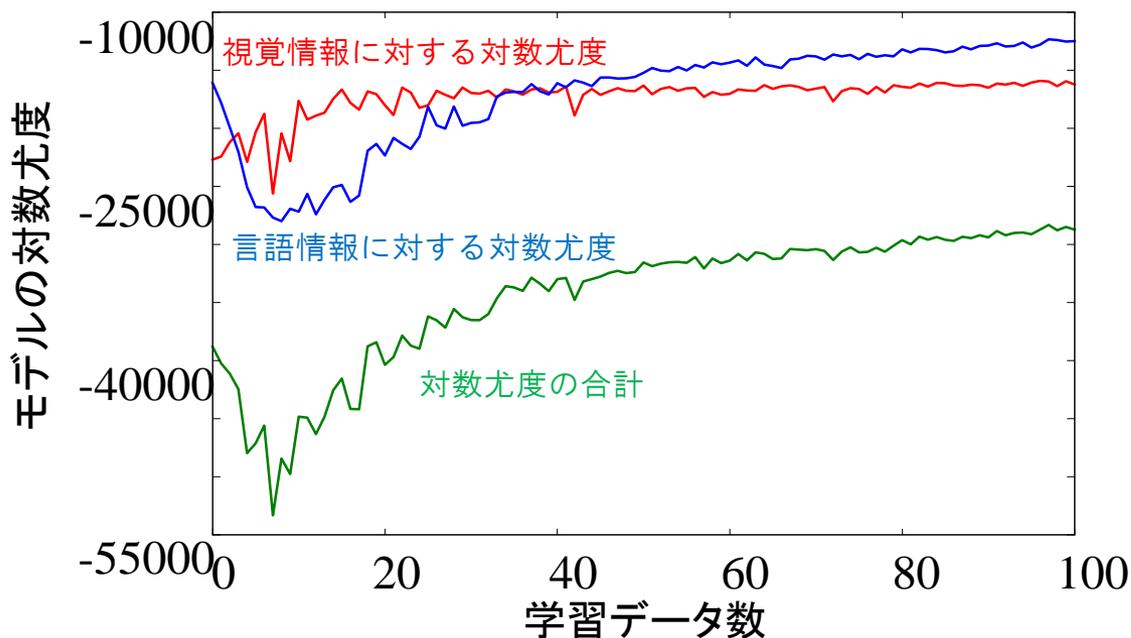


- ▶ モデル構造学習が発達であり、かつオンライン学習の場合は発達は必要



音声認識性能の向上（発達）に伴う学習の変化

- ▶ 学習物体数に従って、音声認識性能が向上する様子をシミュレーションで再現し学習
- ▶ 提案モデルにおける対数尤度の変化を観測
 - ▶ 対数尤度＝分類する際の重要度



- 音声認識性能の向上に伴い視覚情報重視な分類から、言語情報重視な分類へと変化
- センサは同じでも言語ごとに異なる概念が形成される

自身の能力(発達?)に合わせて学習の仕方を変化させている

まとめ

まとめ

- ▶ **本発表では概念・語彙・語意獲得ロボットを紹介**
 - ▶ マルチモーダル情報から概念を形成
 - ▶ 人の教示発話から言語モデルを獲得
 - ▶ 概念と言語モデルが相互に影響することで、双方の精度が向上
- ▶ **このロボットを事例に発達が必要なかを考察**
 - ▶ 学習を繰り返すことで徐々に性能が向上
 - ▶ その過程でモデルが徐々に複雑な構造へと変化
 - ▶ 変化した構造が学習にも影響を与えている
 - ▶ 「発達」の定義次第では、このロボットは発達している、と考えることができる



宣伝

- ▶ ロボット学会(9/7~9@山形大)で2つのオーガナイズドセッションを企画(投稿・申込×切:7/13)
- ▶ 確率ロボティクスとデータ工学ロボティクス

センサデータの収集・管理から, 画像・点群・音声処理や運動計画・行動学習などの要素技術, 記号創発やデータ工学ロボティクスへの展開にいたるまで幅広い分野の研究発表を募集いたします。

- ▶ 人工知能学会との共催セッションなので**ロボット学会非会員でも投稿できます!**(詳しくは中村まで)
- ▶ インテリジェントホームロボティクス

本セッションは, 実際の現場で利用できるロボットソリューションを目指し, 大学・研究機関・企業・現場が一体となって, 1)現場で動くホームロボットに要求されるサービスの在り方の検討, 2)技術の体系化とソリューションのパッケージ化, 3)フィールド評価実験の標準化, について議論する。

是非, ご投稿ください!