

# “意味”を伝達するセンサネットワークの提案

高田 諭† 橋本周司‡

早稲田大学 応用物理学科†‡

## 1. はじめに

人とロボットのコミュニケーションにおいて、ロボットが人の提示する“意味”を理解することができれば、環境に依存しない柔軟なコミュニケーションが可能になると考えられる[1]。そこで、我々は人の提示する情報や物体に共通して付加されている“意味”を読み取り出力するセンサ「意味センサ」を提案し、画像情報を元にその有効性を検証してきた[2]。本稿では、これに音声情報を基にした意味センサを組み合わせたセンサネットワークを提案し、より環境に依存しないロバストな意味の理解が可能であることを示す。

## 2. 提案手法

従来型センサの出力はセンサ毎に出力形式が異なるローデータであるため、各センサ専用の出力機器や処理プログラムを用意する必要があった。例えば、音声認識に画像処理を付加して耐雑音性を向上しようとするには画像処理用のハードウェアとソフトウェアが必要となるなど、柔軟に対応するのが難しい。そこで我々は、意味のカテゴリをあらかじめ定めて、センサ信号がどのカテゴリに当てはまるかを判定し、意味として出力する意味センサ(図1)のネットワークを検討している。意味を仲介とすることにより、センサが変わってもロボット側の変更は不要なばかりでなく、新たな機能の追加も容易になり、柔軟なシステムを構成することができると思われる(図2)。

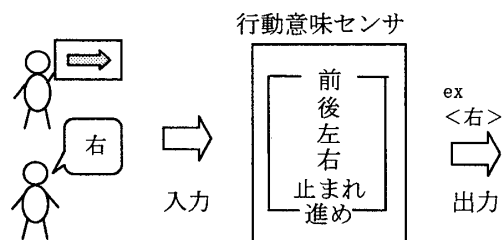


図1 意味センサの定義

A proposal of a sensor-network to communicate with meaning  
†Satoshi Takada, ‡Shuji Hashimoto · Department of Applied  
Physics, Waseda University

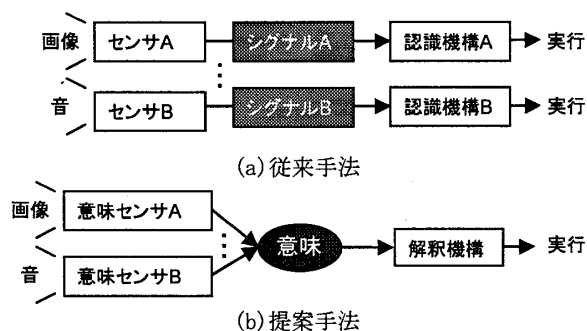


図2 従来手法と提案手法の比較

表1 発話に対する意味の出力

発話(入力)	意味(出力)
まえ, front	前
うしろ, back	後
ひだり, left	左
みぎ, right	右
とまれ, しゆりよう, stop	止まれ
すすめ, いけ, go	進め

## 3. 音声情報を基にした行動意味センサ

人がロボットに行動を指示する場合を仮定し、ロボットの基本的な行動に関する意味として動作「止まれ、進め」および移動方向「前、後、左、右」を音声情報から出力する行動意味センサを作成した。表1に発話に対する意味の出力を、図3にはセンサの構成を示す。ここで音声認識モジュールとしてSRM-10P(アルファプロジェクト)を使用し、不特定話者の「まえ」や「front」等の発話に対し「前、後、左、右、止まれ、進め」のいずれかの意味、または認識不能を表す「不明」を出力する。

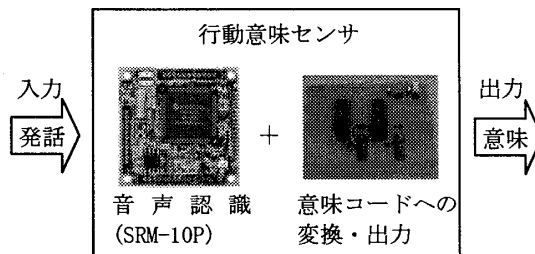


図3 行動意味センサの構成

表 2 意味の理解率[%]

		「前」	「後」	「左」	「右」	「止まれ」	「進め」	平均
(1)画像のみ	外乱光なし	90.0	93.3	83.3	86.6	83.3	93.3	88.3
	外乱光あり	23.3	33.3	23.3	0.0	40.0	30.0	25.0
(2)音声のみ	雑音なし	100	100	90.0	100	100	100	98.3
	雑音1	50.0	26.7	60.0	20.0	40.0	26.7	37.2
	雑音2	13.3	3.3	6.7	10.0	16.7	10.0	10.0
(3)画像と音声	外乱光あり	100	93.3	96.7	96.7	100	100	97.8
	雑音2	90.0	100	83.3	86.7	83.3	100	90.6
	外乱光と雑音2	36.7	50.0	26.7	10.0	46.7	33.3	33.9

#### 4. 意味の出力実験

##### 4. 1 実験条件

画像情報と音声情報の意味センサの組み合わせで、ロボットの動作の信頼性を向上させることが可能であることを確認する実験を行なうために、以下の状況で実験を行った。

##### (1) 画像意味センサのみを用いる場合

外乱光のある場合とない場合で実験を行った。外乱光はアイランプ 300W フラッド(岩崎電気株式会社)を用い、センサに向かって右手に設置した。また太陽光等による色とびを再現するために、センサと提示物体の間から約 15-20cm の距離に配置した。

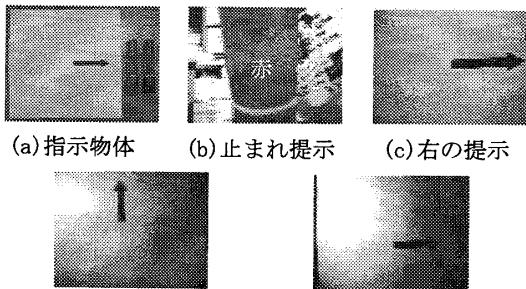
##### (2) 音声意味センサのみを用いる場合

雑音のある場合とない場合を想定した。雑音は Radio Nikkei(NIKKEI RADIO BROADCASTING CORP)を用い、センサから約 30cm はなれた位置にステレオスピーカーを設置した。また、雑音の音量については、小さい場合(55-60[db])を雑音 1、大きい方(69-73[db])を雑音 2 とした。

##### (3) 画像と音声の両方を用いる場合

外乱光のある場合、雑音のある場合と外乱光と雑音両方ある場合の比較実験を行った。

画像センサ用に提示する指示物体は、意味「前、後、左、右、止まれ、進め」に対応するものをあらかじめ決めておき[2]、それぞれ 30 回ずつ提示した。指示物体と提示した様子、そして、外乱光による影響の様子を図 4 に示す。



(a) 指示物体 (b) 止まれ提示 (c) 右の提示  
(d) 外乱光あり, "前" の提示 (e) 外乱光あり, "右" の提示  
図 4 提示物体と提示例の様子

##### 4. 2 結果と考察

音声・画像の 2 種の意味センサの出力が、矛盾するか共に「不明」の場合は、決定不能として、意味解釈を行なった。それぞれの意味・環境において指示を正しく理解できた割合[%]を表 2 に示す。画像と音声を組み合わせることで外乱光のみの場合 25.0%から 97.8%へ、雑音 2 のみの場合には 10.0%から 90.6%へと改善され、外乱光や雑音のない場合と同等に意味を理解することができた。外乱光と雑音の両方がある場合(33.9%)は、画像あるいは音声単独の場合に比べて多少の向上は見られたが、安定した意味の理解は難しかった。表 2 で、外乱光がある場合「右」の理解率が比較的に低いのは、外乱光の設定によるものである(図 4 (e))。設定によっては他の意味の理解率が低下することもある。

#### 5. まとめ

音声情報を基にした意味センサを作成し、画像情報の意味センサと組み合わせたセンサネットワークを検討した。意味を共通の情報としてセンサネットワークを構成するため、異なる観点の意味センサを追加することにより、大幅なシステムの更新を行なわなくても指示や状況の理解をさらに強力にできる。現在、センサの種類を増やしてロボットに実装することを検討している。

本研究の一部は、早稲田大学特定課題(課題番号: 2007B-143) および早稲田大学理工学研究所プロジェクト研究の「ロボット技術の社会適用に関する研究」として行われた。

#### 参考文献

- [1] 田中秀幸, 矢入健久, 町田和雄: "意味的情報を扱う自律ロボットシステム", 人工知能学会主催 MYCOM2006, セッション 4-1, 2006.
- [2] 高田諭, 酒井幸仁, 橋本周司, "意味をインタフェースとしたヒューマン-ロボットコミュニケーションのための意味センサの提案," ViEW2007 講演論文集, pp.123-128, 2007.