

音環境理解研究からのロボット聴覚の構築

奥乃 博

(京都大学・大学院情報学研究科・教授)

【研究の概要等】

ロボット自身に装着したマイクロフォン（耳）で聞くという「ロボット聴覚」は、ロボットを実世界で配備するために不可欠な機能である。ロボットが実世界で聞く音は、複数の音源からの音が混ざった混合音があることが多い。音環境理解は混合音から音源定位・音源分離・分離音認識を通じて、音による環境理解を行う研究である。私たちはこれまでに進めてきた音環境理解研究の成果を、ロボット聴覚に応用し、聖徳太子のように複数の音を聞き分け、理解するような聴覚機能、すなわち、「聖徳太子ロボット」の実現を行う。具体的な研究項目は以下の通りである：

- (1) 個別ロボットにも特定動作環境にも依存しない「事前知識最少の」汎用技術を開発。
- (2) 階層的な視聴覚情報統合による複数移動話者の定位・分離・認識技術を確立。
- (3) ロボットの身体を活用したシンボル獲得とコミュニケーション機能を実現。
- (4) 音声だけでなく、音楽や環境音の認識を用いた複数話者との対話技術の確立。

【当該研究から期待される成果】

次のような技術が確立でき、ロボットと人との共生がより知的なレベルで可能となる：

- (1) 複数話者同時発話を認識する聖徳太子ロボットがどのロボットにも手軽に実現。
- (2) 音声・音楽・環境音を聞き分け、認識する機能を実現。
- (3) 音響信号記号変換によるこれまで未使用だった音情報の活用。
- (4) ロボット自身が発生するモータ雑音等の自己モデルを獲得し、その音を抑制して聴きたい音を強調する機能の実現。

【当該研究課題と関連の深い論文・著書】

- ・ D.F. Rosenthal and H.G. Okuno (eds.): *Computational Auditory Scene Analysis*, p.399+xiii, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1998.
- ・ 奥乃博: 音環境理解から見たロボット聴覚、*日本音響学会誌*、63:1(Jan. 2007) 29-34.
- ・ H.G. Okuno, K. Nakadai, *et al*: Human-Robot Non-Verbal Interaction Empowered by Real-Time Auditory and Visual Multiple-Talker Tracking, *Advanced Robotics*, 17:2(2003) 115-130

【研究期間】 平成19年度－23年度

【研究経費】 47,500,000 円
(19年度直接経費)

【ホームページアドレス】 <http://winnie.kuis.kyoto-u.ac.jp>