

## 音環境理解に基づくロボット聴覚の構築

Development of Robot Audition based on  
Computational Auditory Scene Analysis



奥乃 博 (OKUNO HIROSHI G.)

京都大学・大学院情報学研究科・教授

### 研究の概要

混合音入力からの音源定位・音源分離・分離音認識という音環境理解の成果を、「ロボット自身の耳で聞く」というロボット聴覚に応用し、複数の音を聞き分け、理解できる汎用的聴覚機能を開発し、「聖徳太子ロボット」を実現する。さらに、ロボット聴覚をオープンソース化し、世界中のロボットを聖徳太子化させて、人とロボットとの共生を促進する。

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：ロボット聴覚，聖徳太子ロボット，音源定位，音源分離，音声認識

### 1. 研究開始当初の背景

- (1) ロボット自身の耳で聞くロボット聴覚は我々が2001年に提唱。音から音響事象を認識・理解する音環境理解研究から、混合音中の音源定位、分離、認識に取り組んできた。
- (2) 当時の大部分の研究は音源定位に留まっており、分離音の音声認識まで踏み込んでいたものは我々ともう1件程度であった。
- (3) いずれも個別ロボットに特化しており、汎用性という視点が欠如していた。

### 2. 研究の目的

- (1) 複数の同時発話を聞き分け、裁いたという聖徳太子のような聴覚機能を実現。
- (2) 音環境理解の3要素技術の適用限界を明確にしたロボット聴覚設計原理の確立。
- (3) 階層的視聴覚情報統合による複数移動話者の定位・分離・認識技術の確立。
- (4) 身体性に基づいた記号獲得と記号接地によるコミュニケーション機能実現法の確立。
- (5) 音を用いた複数話者との対話技術の確立。

### 3. 研究の方法

- (0) ロボット聴覚テストベッド用ヒューマノイドの開発: HRP-2W への視聴覚センサの搭載。
- (1) 音環境理解技術の汎用化。
- (2) 視聴覚情報統合による音環境理解技術のロボスタ化。
- (3) RNNによる身体性を考慮した記号獲得。
- (4) バージン発話が許容される複数話者との対話システムの開発。
- (5) 音環境理解技術の出力に対する可視化技術の開発。

### 4. これまでの成果

### (1) ロボット聴覚ソフトウェア HARK の構築とオープンソースとしての公開と講習会実施。

2008年に初版を公開し、京都とソウルで講習会実施。2009年バージョンアップし、横浜(写真)とパリで講習会実施。いずれの講習会で



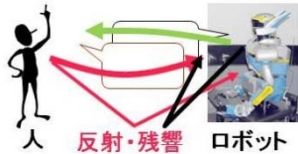
もロボットの頭部を持ち込み3話者同時発話認識のライブデモを実施。

ロボットの身体性を考慮した、すなわち、身体のインパルス応答を活用して、MUSIC法により音源定位性能の向上を、同様にGHDSS法により音源分離性能の向上を、また、ミッシングフィーチャマスク自動生成により音声認識性能の向上を図っている。また、ロボット自身のモータの動作音を方向情報でキャンセルする単純な自己生成雑音抑制機能も有する。この結果、三話者同時発話認識、あるいは、方向性のある雑音下での音声認識がロボスタとなった。[文献1]

### (2) ロボットの頭部の動きを組み合わせた音源定位での前後の曖昧性の解消。

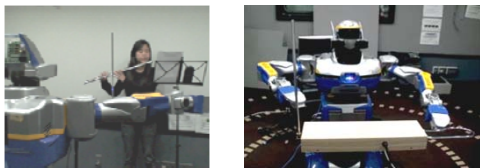
ロボットに2つのマイクしかない場合、音源定位では前後どちらに音源があるのか分からないことがある。ロボットが頭を、水平方向だけでなく、最初に少し傾きを入れてから動かすと、人と同様に曖昧性がよく解消できることが判明した。[文献4]

### (3) セミブラインドICA(独立成分分析)による残響環境下での自己生成音抑制機能。



自然なインタラクションを行うには、ロボットの発話中であっても人が自由に割り込んで発話できた方がよい。このようなバージョン発話は、双方向通信が可能な電話では日常的に使われているが、対面での会話では難しい。ロボット自身が自分の発話すべき信号が分かっているため、自分の耳に聞こえる、反射や残響が入った自分の声を抑制し、相手の声だけを残響を除いて分離するセミブラインド ICA による自己生成音抑制機能を実現。これにより、バージョン発話を許容する音声対話システムの構築が可能となった。

(4) **ロボット聴覚応用としての音楽ロボット。** 上記自己生成音抑制機能を利用し、ロボットの歌声や演奏音を抑制し、相手の演奏音だけを聞き分けることにより、ビート追跡とテンポ予測の性能が向上し、アンサンブル演奏を指向した音楽ロボットの開発が可能となった。現在、電子楽器テルミン演奏ロボットを汎用的アーキテクチャで構築。さらに、HRP-2, HIRO, ASIMO に実装し、有効性を確認した。



#### (5) HARK 出力の 2D 空間での可視化

Schneiderman が提唱した “Overview-First, Zoom and Filter, then Details on demand” に基づく音環境提示法を開発し、オンライン及びオフラインでの表示検索を可能にした。

#### (6) 米国ロボットベンチャ Willow Garage 社のテレプレゼンスロボットへの HARK 搭載。

HARK はこれまでに、HRP-2, SIG2, Robovie, ASIMO 等に搭載されてきた。今回他流試合として Willow Garage 社の Texai (下図中央) に HARK を搭載し、前項の可視化システムを応用して、遠隔ユーザが指定した範囲の音だけを提示するカクテルパーティ効果の機能を実現した。事前準備もほとんどせずに、音環境理解に対するテレプレゼンスの概念を短期間でまとめ、実装できたことは重要な経験となった。



#### 5. 今後の計画

(1) 音量差のある複数音源に対する音環境理解技術の汎用化。

(2) 顔の動きにロバストな視聴覚情報処理による音環境理解技術の開発。

(3) MT (Multiple Timeline) RNN による身体性による記号獲得の開発。

(4) ロボット聴覚を応用した音楽ロボットと人とのアンサンブル演奏への展開

(5) バージン発話を許容した複数話者との対話システムの洗練化。

(6) 音環境理解技術を用いた可視化技術の応用。

(7) 聖徳太子ロボットとしての観点から評価を行う。

#### 6. これまでの発表論文等

1) K. Nakadai, H. G. Okuno, H. Nakajima, Y. Hasegawa, and H. Tsujino: Design and Implementation of Robot Audition System "HARK", *Advanced Robotics, in print*, 2010, VSP and RSJ.

2) 武田龍、中臺一博、駒谷和範、尾形哲也、奥乃博: 残響下でのバージョン発話認識のための多入力独立成分分析を応用したロボット聴覚、*日本ロボット学会誌*, Vol. 27, No. 7 (2009) 782-792.

3) 神田尚、尾形哲也、駒谷和範、奥乃博: 人工神経回路モデルと声道物理モデルを用いた母音模倣モデルに基づく音素獲得シミュレーション、*日本ロボット学会誌*, Vol. 27, No. 7 (2009) 802-813.

4) H-D. Kim, K. Komatani, T. Ogata, H.G. Okuno: Human Tracking System Integrating Sound and Face Localization using EM Algorithm in Real Environments, *Advanced Robotics*, Vol.23, No.6 (May 2009) 629-653,

5) 池田智志、駒谷和範、尾形哲也、奥乃博: マルチドメイン音声対話システムにおけるトピック推定と対話履歴の統合によるドメイン選択手法、*情報処理学会論文誌*, Vol. 50, No.2 (Feb., 2009) 488-500.

6) T. Otsuka, K. Nakadai, T. Takahashi, K. Komatani, T. Ogata, Hiroshi G. Okuno: Design and Implementation of Two-level Synchronization for Interactive Music Robot, *Proc. of the 24th AAAI Conference on Artificial Intelligence, accepted*, 2010.

7) T. Mizumoto, R. Takeda, K. Yoshii, K. Komatani, T. Ogata, H.G. Okuno: A Robot Listens to Music and Counts Its Beats Aloud by Separating Music from Counting Voice, *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1538-1543, IEEE, 2008. **Award for Entertainment Robots and Systems (NTF Award) Nomination Finalist.**

ホームページ等

<http://winnie.kuis.kyoto-u.ac.jp/HARK/>

<http://winnie.kuis.kyoto-u.ac.jp/SIG/>