

研究種目：基盤研究（S）

研究期間：2007～2012

課題番号：19100003

研究課題名（和文） 音環境理解に基づくロボット聴覚の構築

研究課題名（英文） Development of Robot Audition based on  
Computational Auditory Scene Analysis

研究代表者

奥乃 博（OKUNO HIROSHI G.）

京都大学・大学院情報学研究科・教授

研究者番号：60318201

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. 研究計画の概要

ロボット聴覚の最下位の機能とは「ロボット自身の耳」から入力される音の処理である。

(1) 複数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するために、音環境理解の3要素技術である音源定位、音源分離、分離音認識の適用限界を明確にした汎用のロボット聴覚設計原理を確立。また、階層的視聴覚情報統合による複数移動話者の定位・分離・認識技術を確立する。

(2) 人の音環境理解及び聴覚メカニズムを解明するために、音声模倣を規範とした音素獲得過程を特に、音素集合を与えない連続音響信号からの音素獲得のモデル化を行う。

(3) ヒューマノイドロボットにおける音声対話機能、特に、ユーザビリティの高い音声対話を実現するには、ユーザが自由なタイミングで割り込んで（バージイン）発話でき、かつ自由な表現で発話できる枠組みが不可欠である。前者は自己発話抑制機能による枠組みを、後者はシステムが解釈できない想定外発話への対処法を開発する。

### 2. 研究の進捗状況

(1) ロボットの体のインパルス応答を測定し、音源定位、音源分離に活用した処理技法を開発。また、方向性雑音抑制機能も追加し、ロボット聴覚ソフトウェア HARK をミドルウェア FlowDesigner 上で統合し、高速化。さらに、音声認識用特徴量の信頼度を2値から連続値に変更し、信頼度計算法を考案し、3話者同時発話の音声認識精度が10%程度向上。

(2) ロボット聴覚テストベッド用ヒューマノイドとして HRP-2W を導入し、HARK を搭載し、複数人とのインタラクションの可能性を検討。また、HARK を SIG2, Robovie, Honda ASIMO に搭載し、その移殖性の高さを確認。さらに、米国ベンチャ Willow Garage 社から招待され、同社テレプレゼンスロボット Texai に移殖し、可視化技術と組み合わせ、その有効性を実証。

(3) 音声模倣による音素獲得過程を、音声バブリング学習と連続音声模倣追加学習という2段階でモデル化。具体的には、声道物理モデル Maeda モデルを用いた身体構造を考慮したランダムバブリングを再現し、神経回路モデル RNNPB を用いた予測エラーに基づく時系列ダイナミクス追加学習を実現。認知発達過程における音素獲得過程が再現でき、さらに、バブリング動作の特性を明らかにした。

(4) バージイン発話に対しては、自己生成音を既知とした独立成分分析（ICA）に基づくセミブラインド音源分離を利用し、自己発話を抑制し、さらにユーザの発話タイミングを利用して解釈を行う手法を開発。これにより、実環境下で音声認識精度が低くても、ユーザ意図の推定が可能となる。ユーザの想定外発話に対しては、ユーザ発話や発話履歴に基づいたヘルプメッセージ選択法を開発。これにより、初めての一般ユーザでもシステムを使いながら使用法を習熟できるようになる。

(5) 自己生成音抑制機能を音楽ロボットに応用し、自分の歌声、演奏音（電子楽器テルミンを演奏）を抑制し、伴奏音のテンポを実時間で追跡し、合奏の性能が向上。

(6) HARK の出力を “Overview first, zoom and filter, then details on demand” という設計指針 (Schneiderman による) に基づき表示する 3D 音環境可視化技術を開発. 定位情報や分離音の実時間提示, 音声認識結果のカラオケ風表示など, 音響版ライフログの提示法を提案. 本表示法を, アーカイブテキストと組合せ, 議論の流れも表示にも応用.

(7) 人の耳のように 2 本のマイクロフォンしかない場合には音源定位で前後に曖昧性が残る問題に対して, 頭を水平に動かすだけでなく, 最初に頷きを入れた方が前後判定が容易になることを実証し, 人の頭の動かし方での知見と一致することを確認. また, HARK から得られる音響特徴量と画像処理から得られる唇特徴量を用いて, 発話区間検出と音声認識の 2 階層情報統合により雑音下, 劣照明下での音声認識の頑健化にも取り組んだ.

### 3. 現在までの達成度

② おおむね順調に進展している.

(理由)

- ① ロボット聴覚ソフトウェア HARK を洗練化し, オープンソースとして公開. さらに, 講習会を国内外が 2 回ずつ実施.
- ② 米国 Willow Garage 社でテレプレゼンスロボット Texai に HARK を実装し, GUI 等の開発も含めて 1 週間で完了. HARK の音源定位・追跡, 音源分離機能の有効性を同社食堂・講堂で実証.
- ③ 自己生成音抑制機能を開発し, 音声対話や音楽ロボットに応用し, 有効性を確認.
- ④ ロボット聴覚小特集を日本ロボット学会誌で組み, 当該分野の普及展開を行った.

### 4. 今後の研究の推進方策

2010 年度は, ロボット聴覚システムの評価に基づく個別技術の再構築に取り組み, 最終年度の 2011 年度は, 総合的評価を行う.

#### 【要素技術の再構築】

- ① 音量差のある複数音源に対する音環境理解技術のロバスト化, 高性能化.
- ② 顔の動きにロバストな視聴覚情報統合による音環境理解技術の開発.
- ③ MTRNN による身体性による記号獲得.

#### 【システムへの展開】

- ④ バージン発話を許容した複数話者との音声対話システムの洗練化.
- ⑤ 音環境理解技術を用いた可視化技術の洗練化と議事録文書提示への応用.
- ⑥ 聖徳太子ロボットの観点からの評価.

### 5. 代表的な研究成果

(研究代表者, 研究分担者及び連携研究者には下線)

[雑誌論文] (計 5 5 件 **【査読有】** + 2 0 件)

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. 査読有

② 武田龍, 中臺一博, 駒谷和範, 尾形哲也, 奥乃博: 残響下でのバズ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*, **27:7** (2009) 782-792. 査読有

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*, **27:7** (2009) 802-813. . 査読有

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*, **23:6** (2009) 629-653, 査読有

[学会発表] (計 7 0 件 **【査読有国際会議のみ】**)

⑤ 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 Conf. on Intelligent Robots & Systems*, 1538-1543, 2008. **NTF Award Finalist**.

[図書] (計 1 件)

[産業財産権]

○出願状況 (計 1 5 件)

名称: 音声認識装置及び音声認識装置のマスク生成法

発明者: 中臺一博, 高橋徹, 奥乃博

権利者: 本田技研工業株式会社

種類: 特許出願

番号: 特願 2009-185164 号

出願年月日: 2009 年 8 月 7 日

国内外の別: 国内

○取得状況 (計 3 件)

名称: 音響信号処理方式, 音響信号処理装置, 音響信号処理システム及びコンピュータプログラム

発明者: 後藤真孝, 吉井和佳, 奥乃博

権利者: 京都大学・産総研

種類: 特許

番号: 第 4318119 号

取得年月日: 平成 21 年 6 月 5 日

国内外の別: 国内

[その他]

ホームページ

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

ロボット聴覚ソフトウェア HARK

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