

【ICT イノベーション創出型研究開発(ユニバーサル・コミュニケーション技術)】

**音情景分解に基づく音メディア・ユニバーサル・
コミュニケーション・システムの研究開発**

(課題番号081707005, 研究期間H20～H22年度)

**猿渡 洋, 鹿野清宏, 川波弘道(奈良先端大)
細井裕司(奈良県立医科大)**

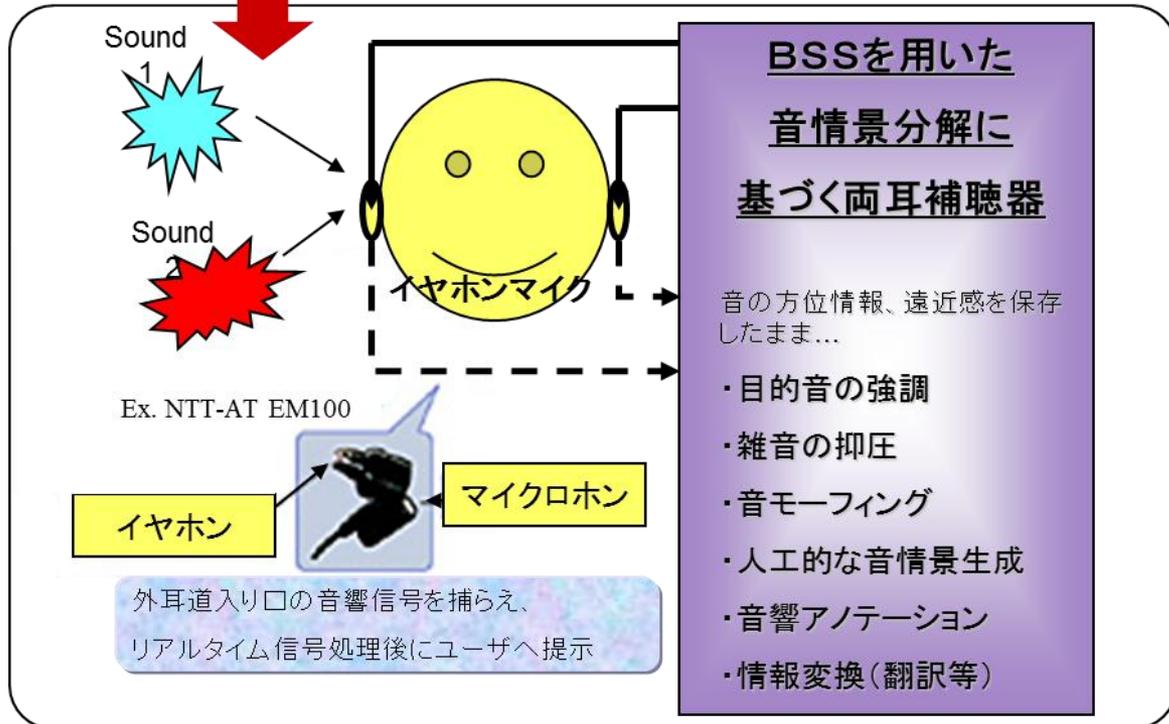
(2011年10月4日)

本研究の概要

- 研究目的: 音情景分解・拡張・合成に基づく新しい音メディア・コミュニケーション・システムを創出する。

※音情景: 音源そのものだけでなくそれを取り巻く音響環境(空間的広がり感など)も含めた単位

- 研究概要: (a) ブラインド音源分離(BSS)に基づく高精度かつ高速な音情景分解
(b) 音情景単位における空間情報の拡張・合成(音のレタッチ)
(c) 上記(a)(b)を統合した両耳補聴器によるコミュニケーション能力拡大



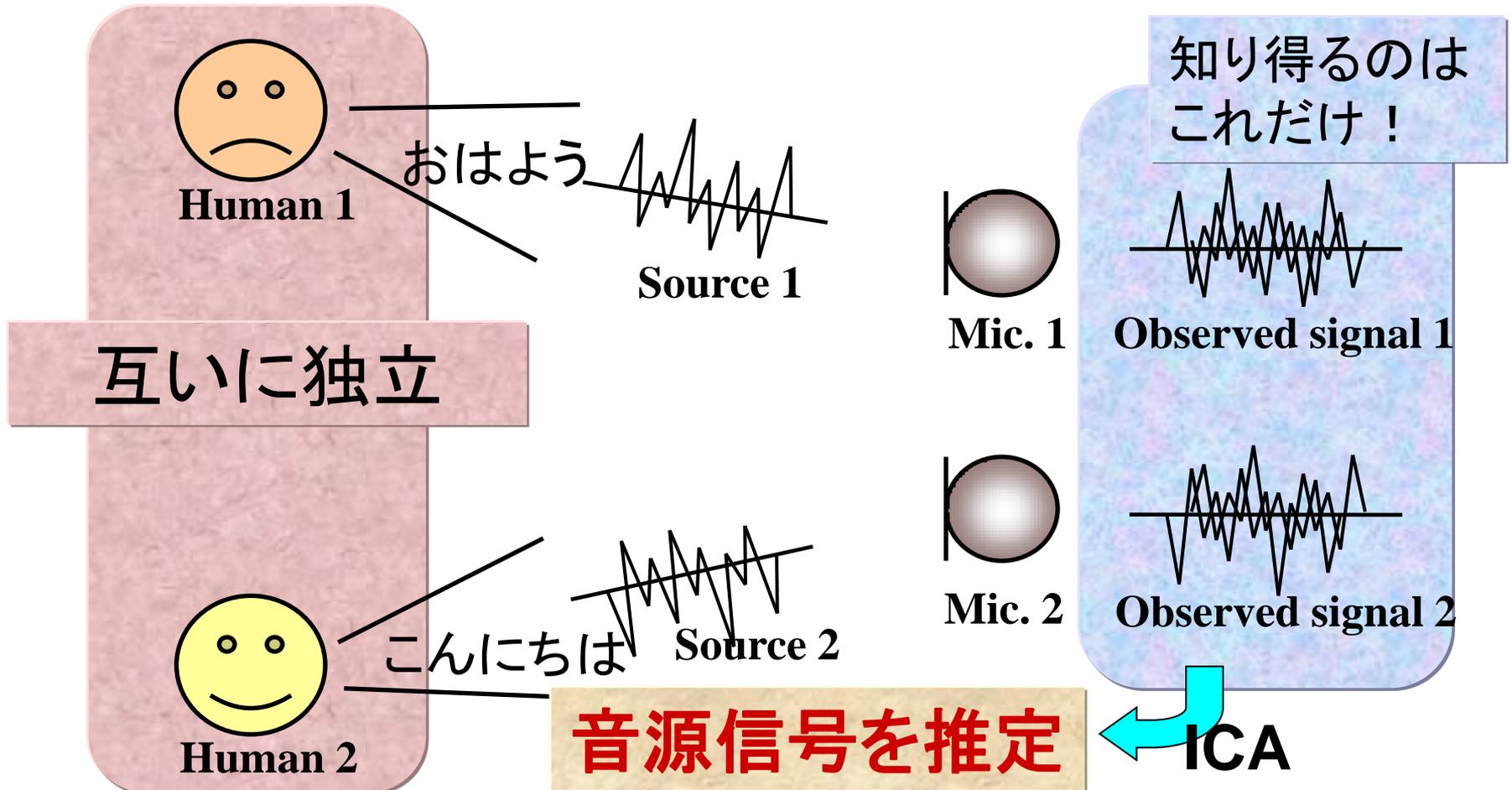
- 研究成果およびインパクト

- ・リアルタイムBSSに基づく新しい音メディア・コンテンツ処理が創出できた。
- ・聴覚的弱者の能力をサポートするユニバーサル・コミュニケーション手段が実装できた。
- ・健聴者の音メディア・コミュニケーション能力も拡大する音ARが実現できた。

基盤要素技術：ICAに基づくBSS

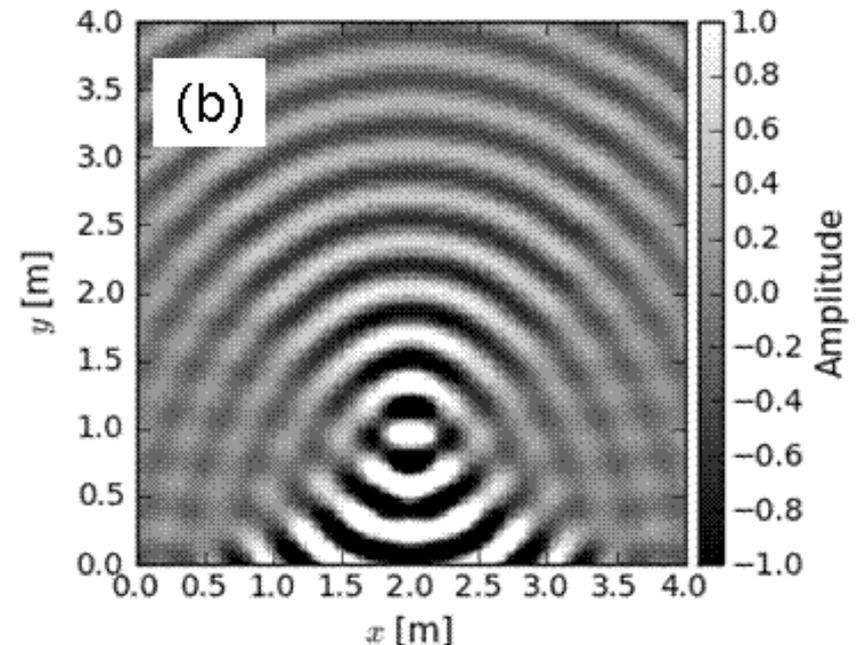
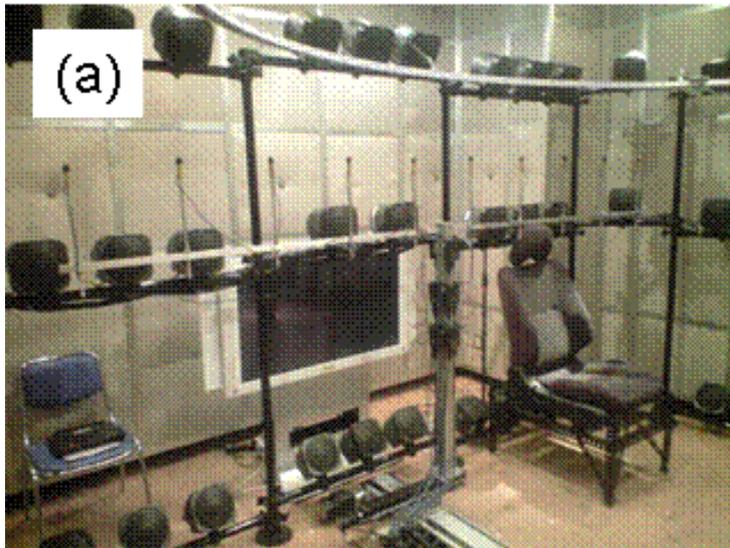
●ICA(独立成分分析)とは:

- 複数の信号が混ざって観測された信号のみから、各音源の独立性(関連の無さ)のみを頼りにして、最適な分離フィルタ行列を推定する技術。
- 音源の種類・方位、マイクの位置・固有特性・誤差などの情報を必要としない。



成果1. 音情景シミュレータの開発

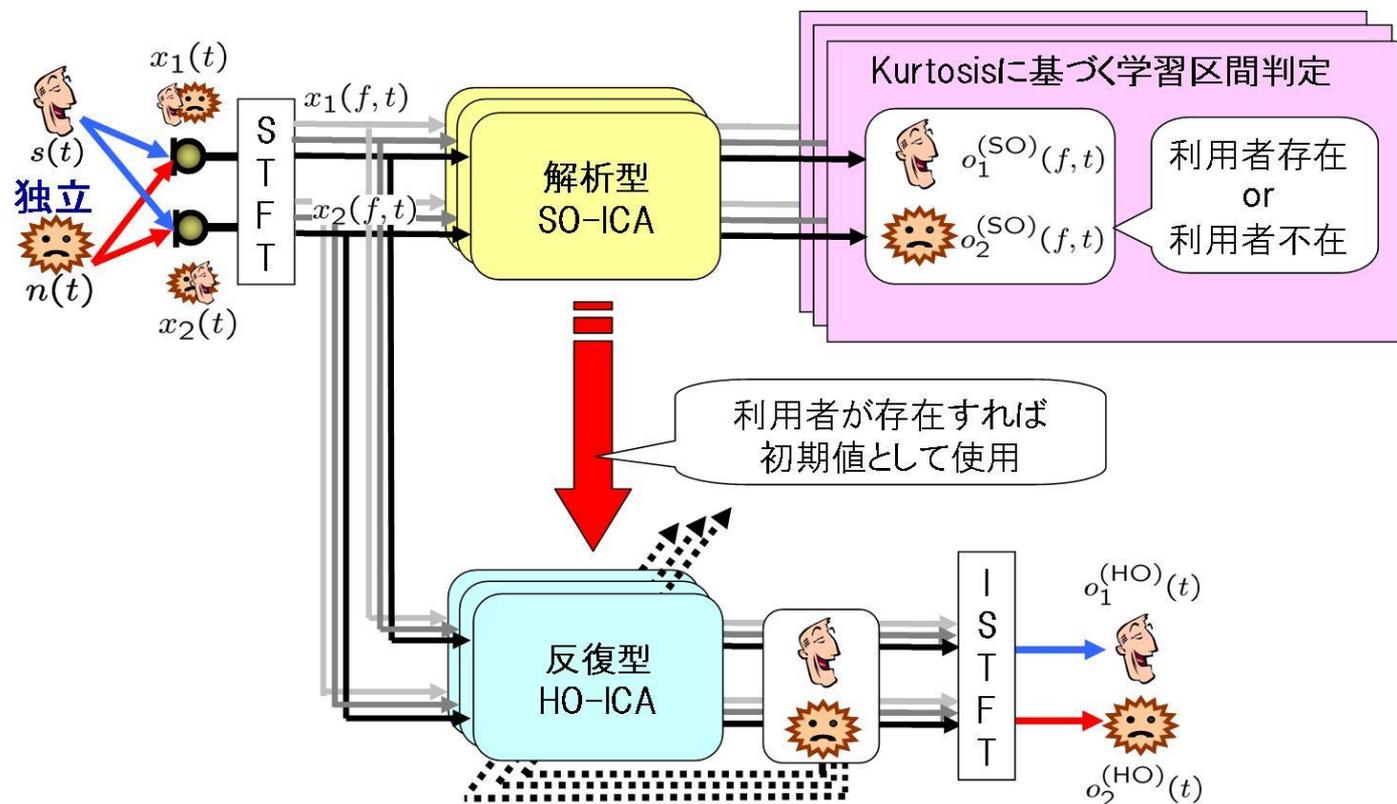
- 音情景データベースの整備に向けて、典型的な人間頭部モデルの全方位インパルス応答(のべ20人分)や多種の環境雑音の収録を行い、本研究に必要な基礎データベースの構築およびそれらを用いたシミュレータシステムを完成した。



(a) シミュレータ外景、(b) 波面合成法による仮想音源の再現例

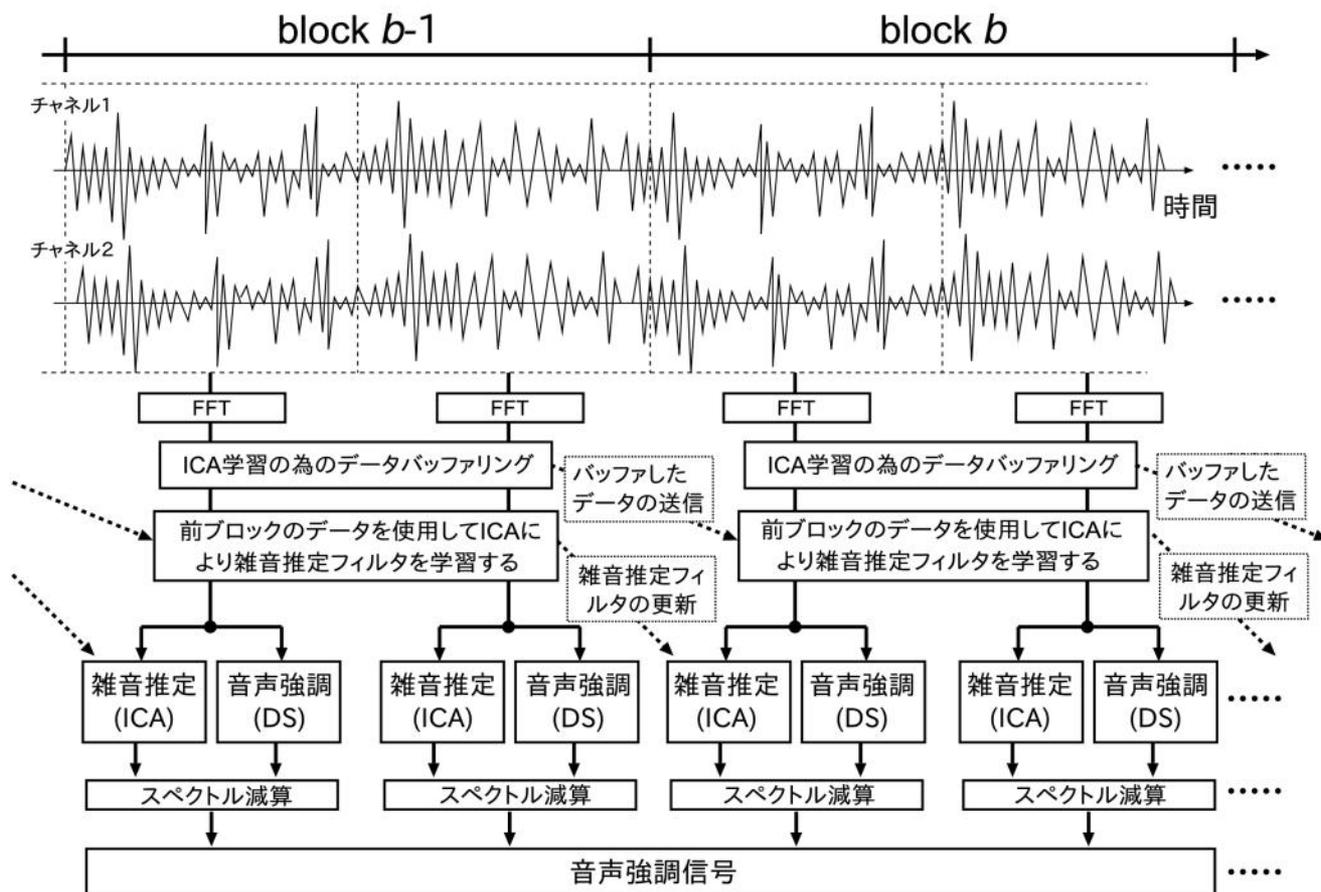
成果2. 高速収束音情景分解アルゴリズム

- ICAの収束速度を向上させるため、反復学習の必要がない解析型二次統計量ICAを併用した発話検出・初期化手法を提案し、**30回**の反復で十分な分離精度が得られることを実証した。



成果3. リアルタイム音情景分解システム[1/3]

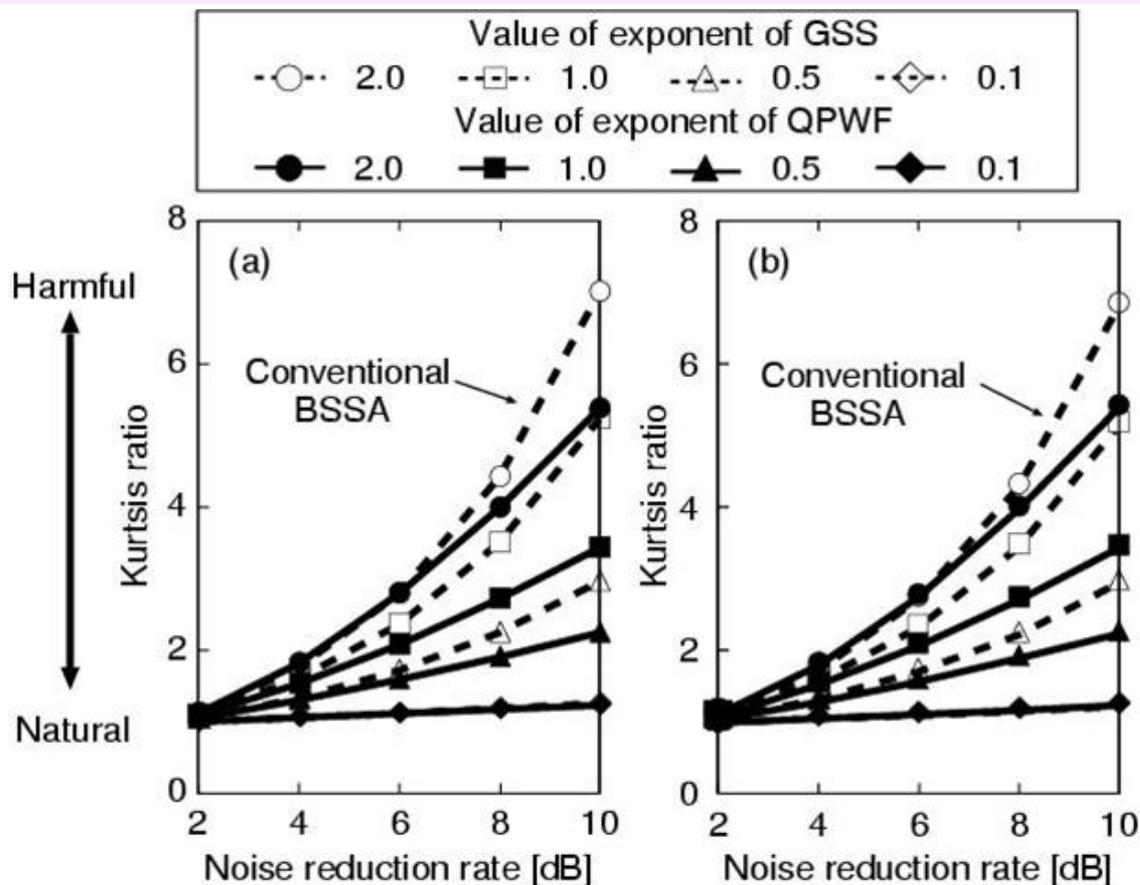
- リアルタイム(入力後100 ms以下で出力)で動作するBSSの開発・実装を行った。これは通常のPC上で稼働するものであり、**実測遅延は56.5 ms**であった。



成果3. リアルタイム音情景分解システム[2/3]

- ポストフィルタの改善を通じて分離性能向上を図った。
- 一般に、従来处理では、「ミュージカルノイズ」と呼ばれる不快な残留雑音が発生する。この発生機構を数理的に解析した結果、信号の確率密度の変形具合と音質との間に相関があることを突き止めた。
- そこで、ICAのポストフィルタのミュージカルノイズ発生量を高次統計量Kurtosis比に基づいて評価し、最適な内部パラメータを設定する枠組みを提案した。これを、一般化スペクトル減算(GSS)及び疑似パラメトリックウィーナフィルタ(QPWF)の2種類に導入した。
- 基礎となる前述リアルタイム音源抽出システムに導入しても、遅延を70 ms以下に抑えることが可能であった。

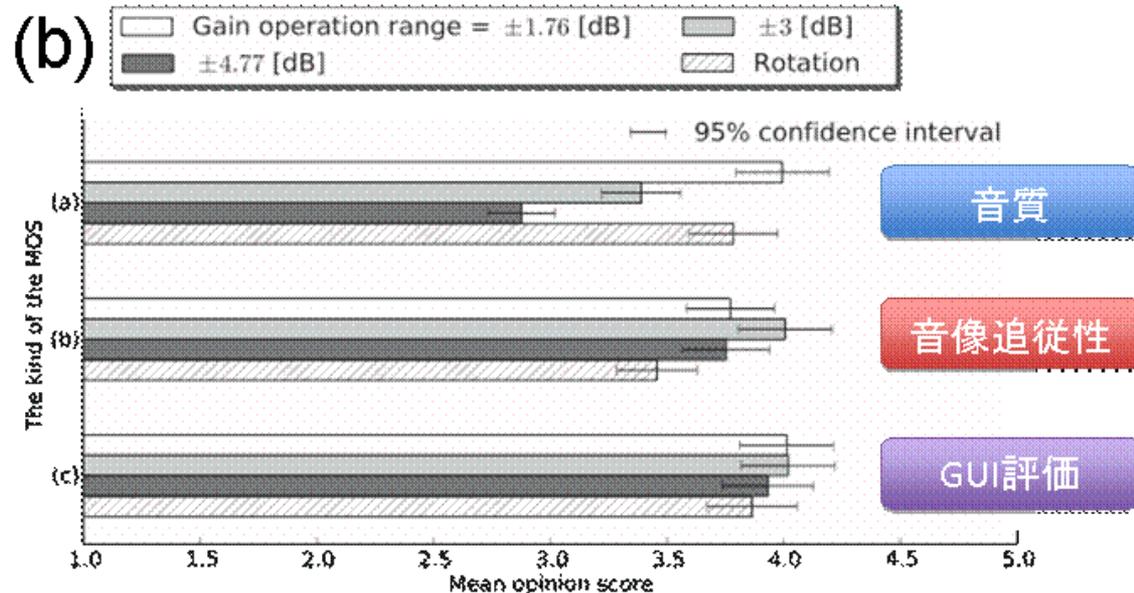
成果3. リアルタイム音情景分解システム[3/3]



従来音源抽出法(BSSA)と改良型ポスト処理(GSS, QPWF)に基づく音源抽出法の客観比較調査結果。(a)は白色雑音、(b)は駅環境雑音の場合の結果である。雑音抑圧量(NRR)は大きいほど、カトシス比は小さいほど品質が良いことを表す。

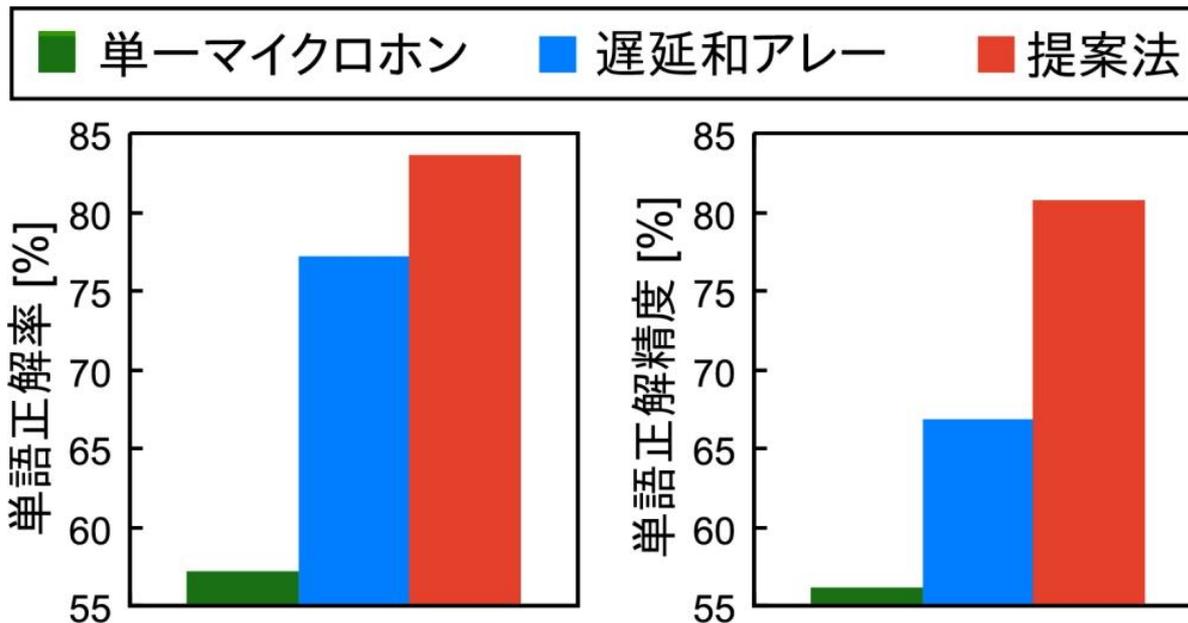
成果4. 音情景特徴量の変換・拡張[1/2]

- 典型的な音情景の拡張処理システムとして、複合音の個別音量操作や音定位の変形を実現する音オブジェクトクラスタリング音量・定位感操作法を提案し、タブレット型PC上にGUIとともに実装しそのユーザビリティ評価を行った。その結果、十分な満足感をもってユーザビリティが実現されていることが示された。



成果4. 音情景特徴量の変換・拡張[2/2]

- 音情景の情報変換として、分離された音情景中から自動的に音声を抽出し内容認識する処理を開発した。
- リアルタイム音情景抽出処理と音声認識システムとを統合し、目的音声の認識精度に関して評価を行った。
駅雑音下にて発声された音声を認識したところ、従来手法と比べて単語認識率の大幅な改善が確認された。



成果5. 音情景分解に基づく両耳補聴器[1/3]

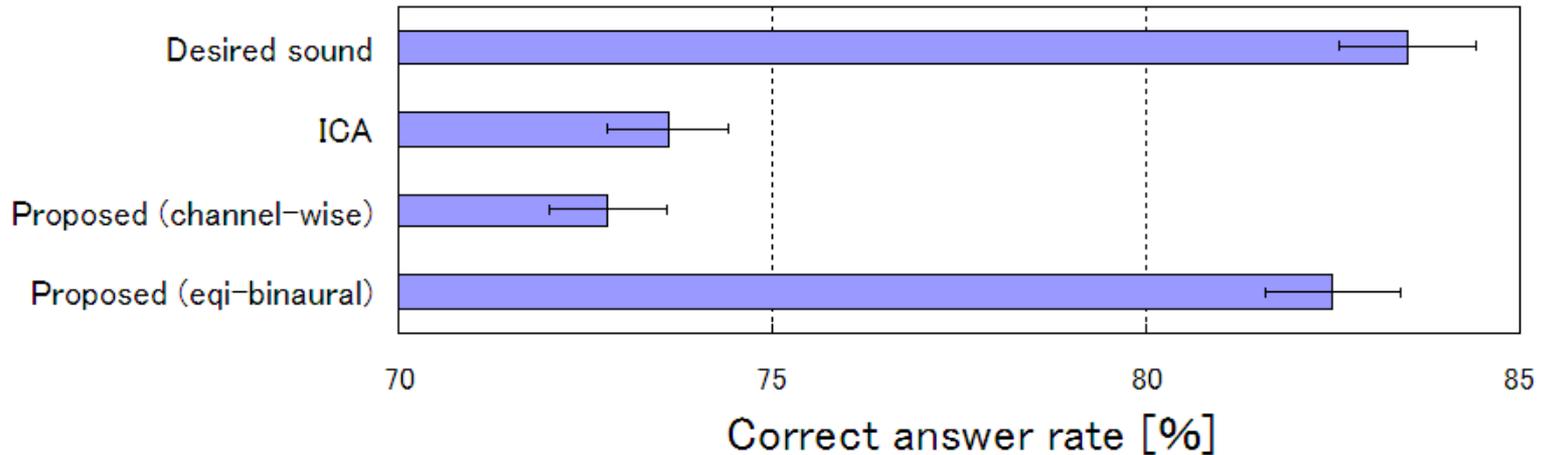
- 音情景分解アルゴリズムを両耳補聴器へ応用するために、ICAによる雑音推定と最小二乗誤差 (MMSE) 規範に基づく振幅スペクトル推定によるポスト処理とを組み合わせ、それに基づく両耳補聴器を提案した。
- ここでは、(1)音像定位保証のため、両耳において同じスペクトルゲインを用いるという制約の元で、(2)MMSE 規範において左右における残留雑音誤差が最小でなければならない、という条件を明らかにした。
- 上記の2要素を満たすため、「両耳共通化最適スペクトルゲイン(equi-binaural)」アルゴリズムを開発した。本手法の特徴は、MMSE規範の元で高いSN比を実現しつつ、人間の音像定位に関して補償を行っている点にある。

成果5. 音情景分解に基づく両耳補聴器[2/3]

- 健聴者による評価実験を行ったところ、雑音抑圧性能は従来法とほぼ同程度であるにもかかわらず、各音情景の音像定位や品質に関して大幅な改善が見られた。

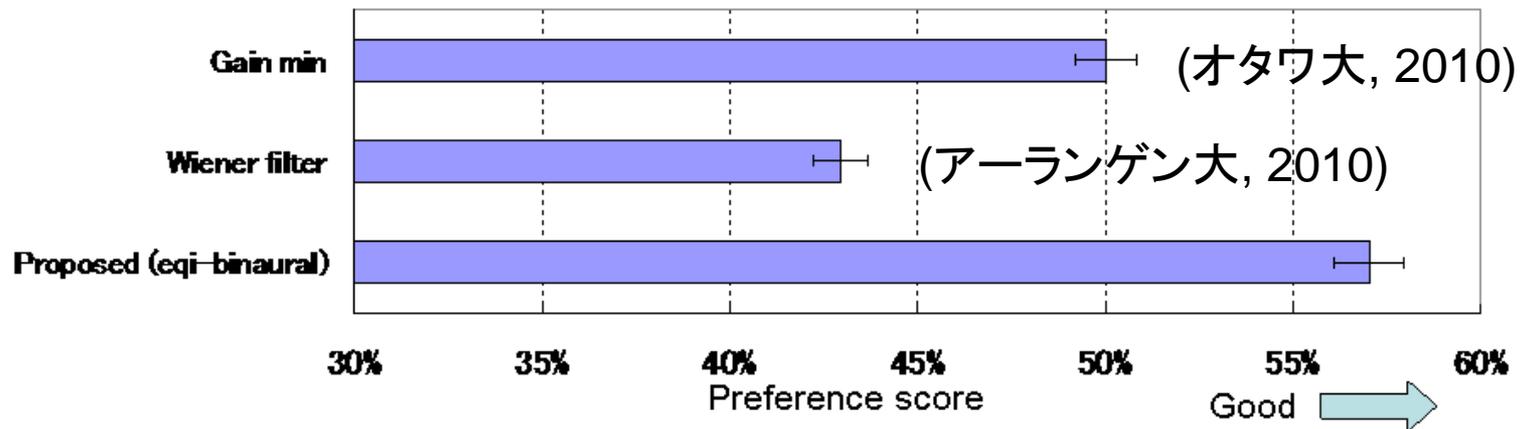
音像定位

評価結果



主観音質

評価結果



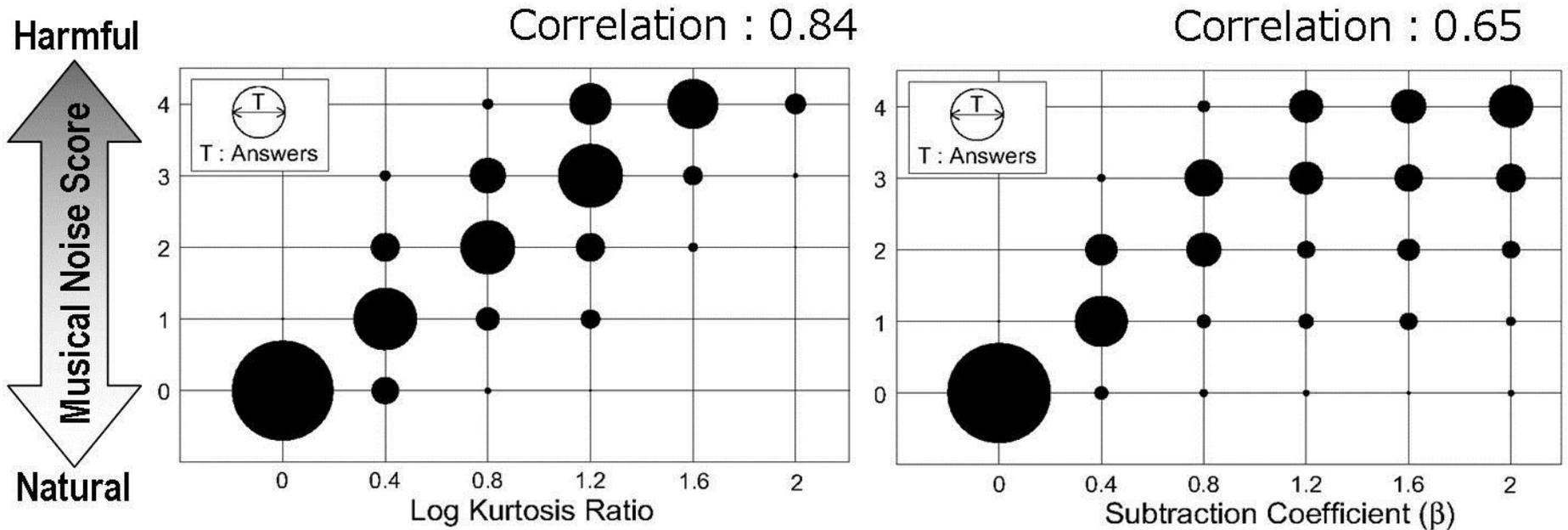
成果5. 音情景分解に基づく両耳補聴器[3/3]

- 両耳補聴器の更なる評価のため、実際の難聴者における語音明瞭度の改善について調べた。
- 奈良県立医大を訪れた重度の感音性難聴者20名(平均年齢69)に対して、57-S語表を用いた片耳による明瞭度試験を行った結果、統計的な有意差を持って音情景分解の効果を確認することは出来なかった。
- 一方、普段から両耳補聴器を使用している重度難聴者を対象にして、普段使用している市販両耳補聴器と提案する音源分離型両耳補聴システムを併用した両耳評価を実施した。被験者が1名なので統計的に有意な結果とは言えないが、約8%の明瞭度改善効果(明瞭度36%から44%の向上)が見受けられ、内省意見として「聞いていて疲れない」というコメントを得た。

当初計画以上の成果：聴感品質の定量化理論[1/2]

- 「音質改善」に関連する研究開発において、「何が人間にとって不愉快な残留雑音か」を解明し、また「その不愉快具合を客観数値で表す」必要性があると考えた。
- これを実現するため、非線形信号処理部におけるミュージカルノイズの発生メカニズムを数理的に解析したところ、信号の確率密度構造の変形具合と音質（雑音の不愉快さ）との間に大きな相関があることを突き止めた。
- よって、音質を客観数値化するために、確率密度構造を特徴付ける高次統計量に基づく「対数Kurtosis比」を提案した。このように、高次統計量と人間の聴覚印象との関係が数理的に解析された例は世界でも初である。

当初計画以上の成果：聴感品質の定量化理論[2/2]



(左) 提案高次統計量Log Kurtosis比と聴覚印象(音色)の関係。
(右) 従来のパラメータと音色の関係。左の方(提案尺度)がより人間の聴覚印象と直線的な関係にあり、高い相関係数が達成されていることがわかる。

まとめ

- 本研究開発では、音情景を基本とする音メディア・コミュニケーション・システムを構築するために、ブラインド音源分離処理を基礎とした音情景抽出・拡張・合成システムを構築した。また、上記システムの実証的アプリケーションとして両耳補聴器システムを構築し、その評価を通じて有効性を確認することができた。
- 本成果は、国際的にもその独創性が認められるに至り、36編の査読付き論文(英文34、和文2編)、61件の口頭発表、5件の受賞を得た。人材育成の面に関しても、多くの修士課程学生(13名)と博士課程学生(3名)が修了および学位取得に至った。以上より、情報通信技術を駆使したコミュニケーション能力拡張およびライフイノベーション創出に大きく貢献できたと言える。