JAIST Repository
https://dspace.jaist.ac.jp/

| | |
|---|---|
| Title | 両耳による選択的聴取を補助する雑音残響環境下音声強調手法の研究 |
| Author(s) | 佐々木，裕吉 |
| Citation | |
| Issue Date | 2012-03 |
| Type | Thesis or Dissertation |
| Text version | author |
| URL | http://hdl.handle.net/10119/10436 |
| Rights | |
| Description | Supervisor: 赤木 正人，情報科学研究科，修士 |

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

# 両耳による選択的聴取を補助する雑音残響環境下音声強調手法の研究

佐々木 裕吉 (0910025)

北陸先端科学技術大学院大学 情報科学研究科

2012 年 1 月 31 日

**キーワード:** 音声強調手法, 両耳による選択的聴取, 雑音残響環境下, Two–Stage Binaural Speech Enhancement with Wiener Filter, Cepstral Mean Subtraction.

音声認識において，雑音や残響の影響を受けることにより，性能の大幅な低下が見られる。また，軽度難聴者は健聴者と比較して，人混みやホールの中など雑音や残響が多い環境下において，聴き取り能力が著しく低下するという報告がある。そのため，雑音や残響を抑圧するために音声強調手法を音声認識などの音声アプリケーションや補聴器に導入する試みが盛んに行われている。これまでに様々な音声強調手法が提案されているが，その中に人の両耳聴機能に着目した手法が存在する。

Usagawa らは Lindemann の両耳聴モデルに基づき，周波数領域両耳聴モデル を提案した。この手法では周波数領域で両耳間位相差と両耳間レベル差を算出する。そして，音源の方向情報の推定を行い，雑音環境下において目的信号を抽出する。処理信号に各音源の方向情報を保存することで，使用者の両耳による選択的聴取を補助することに成功している。Li らによって提案された Two–Stage Binaural Speech Enhancement with Wiener Filter (TS–BASE/WF) では， 1 つの処理体系で雑音抑圧を行わず雑音推定部と雑音抑圧部から成る 2 段階の処理体系を持つことにより，処理性能を向上させている。

音声強調手法の使用環境を考えた場合，屋内での使用も考えられるため，雑音と残響を同時に抑圧する必要がある。室内インパルス応答 (RIR) において，室の大きさに依存した時刻を境界としたとき，初期反射と後部残響に区別することができる。初期反射は壁からの単一な反射音と考えることができるため，目的信号との相関が高くなる。後部残響は複数の反射音が重なることにより目的信号との相関は低いが，部屋中に拡散していることになる。このような特性を持つ残響と雑音を同時に抑圧できる両耳による選択的聴取を補助する音声強調手法はほとんどないと言える。

本研究では，雑音残響環境下において両耳による選択的聴取を補助する音声強調手法の構築を目的とする。周波数領域両耳聴モデルは，音源方向推定にクロススペクトルを用いているため，残響環境下において性能の低下が見られる。一方，残響環境下における

Copyright © 2012 by Sasaki Yuuki

TS–BASE/WF の動作を考えた場合，雑音推定部ではクロススペクトルを用いないため，残響環境下でも目的信号以外の音の推定は可能であると考えられる。しかし，雑音抑圧部では目的信号と雑音が無相関であることが前提となる Wiener Filter を用いているため，処理信号に影響を及ぼすことが予測される。

本論文では，残響環境下における TS–BASE/WF の性能を測定することで，上記の仮説である「TS–BASE/WF の雑音推定部が後部残響の推定に有効であり， Wiener Filter を用いた雑音抑圧部が初期反射に対して性能を発揮できないこと」を示す。これを検証するため，まず残響環境下における TS–BASE/WF の耐性評価実験を行う。そして，初期反射と後部残響が TS–BASE/WF により抑圧可能であるか実験により確かめる。この結果，TS–BASE/WF の雑音抑圧部における Wiener filter では，初期反射の抑圧に効果を示さないことが示された。

次に明らかにされた TS–BASE/WF の問題点から，TS–BASE/WF の改良案を検討した。その結果，Cepstral Mean Subtraction (CMS) を TS–BASE/WF の前処理として採用した。CMS は初期反射の抑圧に効果的である。したがって，TS–BASE/WF の前処理として用いた場合，後部残響成分のみが処理信号に残ることになり，TS–BASE/WF が初期反射の影響を受けることがない。最後に構築された改良手法と TS–BASE/WF の性能比較実験を行った。実験は残響環境下と雑音環境下，そして雑音残響環境下の条件下で行った。その結果，残響環境下と雑音残響環境下，共に改良手法の性能が TS–BASE/WF の性能を上回った。このことから，雑音残響環境下において両耳による選択的聴取を補助する音声強調手法の実現の可能性が示唆された。本研究で提案した TS–BASE/WF の改良手法を音声アプリケーションや補聴器に導入することで，性能の向上が図れる。