

Title	変調伝達関数に基づく骨導音声のブラインド回復法に関する研究
Author(s)	衣笠, 光太
Citation	
Issue Date	2009-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8094
Rights	
Description	Supervisor: 鷗木祐史 准教授, 情報科学研究科, 修士

変調伝達関数に基づく骨導音声のブラインド回復法に関する研究

衣笠 光太 (710021)

北陸先端科学技術大学院大学 情報科学研究科

2009 年 2 月 5 日

キーワード: 変調伝達関数, 骨導音声, フィルタバンク, 明瞭度, ブラインド回復法.

工場や, 作業現場といった高騒音環境下での音声によるコミュニケーションは非常に困難であり, 作業の安全化や効率化のため, 高騒音環境での音声による円滑なコミュニケーションを可能とする方法が求められている. 現在までに, 数多くの方法が提案されているが, 中でも骨伝導を利用した, 骨導マイクを用いて音声を収録する方法は非常に有効な手法である.

骨導マイクは発話の際の顔の振動をピックアップし, 音声を収録するものである. 頭蓋骨と空気のインピーダンスの差が非常に大きいことから, 骨導マイクは外部雑音の影響を受けることなく音声を収録することが可能である. しかし, 骨導マイクを使って収録された音声(骨導音声)は, 顔の骨や皮膚を伝って伝達されることから, 音質が悪く, 音声明瞭度が低い[1]. そのため, 骨導音声を利用して音声コミュニケーションを行うには, 骨導音声の音質や明瞭度を通常我々が耳にする音声(気導音声)と同等に回復する必要がある.

骨導音声は, 高い周波数帯域ほどパワーが減衰することが分かっている. そのため, 骨導音声の音質改善の最も簡単な手法として現在用いられているのが高域強調である. しかし, 骨導音声の減衰作用は, 骨導マイクを設置する位置や, 話者, 発話内容などにより複雑に変化するため[2], 高域強調ではその変化に対応することができない.

この問題に対処した方法として, クロススペクトル法[3]や長時間 Fourier 変換[4]を用いて骨導音声と気導音声の間の伝達特性を求め, その逆特性を利用して音声回復を行う方法がある. また, 伝達特性を適時学習しながら適応フィルタリングにより音声回復を行う手法も提案されている[5]. 逆フィルタリング法と呼ばれるこれらの方法は[3], [4], [5], 骨導音声の周波数成分を回復させるが, 同時にエコーやアーティファクトを生み出してしまいう問題点がある. また, これらの手法は気導音声の情報を必要とするため, ブラインド処理にはなっていない.

一方, 鵜木らは, 気導音声と骨導音声の間の変換関係を伝達特性とみなし, 気導・骨導音声を同時収録した大規模データベースを用いて変換関係の解析を行い, 骨導音声の音

質・明瞭度の改善方法を検討してきた [6], [7]. 彼らの研究のコンセプトは, 音源フィルタモデルを仮定し, 線形予測分析と変調伝達関数 (MTF) の二つの側面から分析を行うことである. これらの解析の結果, 鷗木らは音源信号ではなくフィルタ情報の回復が骨導音声の回復に重要であることを明らかにした. また, これに基づき 2 つの手法が提案されている. 一つは, Vu らが提案した周波数領域での線形予測分析に基づく骨導音声回復法である [6]. この手法はブラインド処理を実現しているが, 気導音声の線形予測係数を学習するプロセスを必要とする. もう一つは, 木村らが提案した時間領域での MTF に基づく骨導音声回復法である [7]. この手法は, 音声明瞭度と関係のある MTF に基づいており, 音声コミュニケーションを行ううえで重要な音声明瞭度を直接回復することができる. しかし, この手法は, 骨導音声回復に気導音声の情報を必要とし, ブラインド処理になっていない. また, 木村らは MTF を表現するための最適なモデルについての考察を行っていない.

骨導音声を利用したコミュニケーションを実現するには, 骨導音声の音声明瞭度の回復が重要な課題である. そのため, 本研究では MTF の概念に着目し, MTF に基づくブラインド骨導音声回復法の提案を目指す.

本研究では, データベースから求めた MTF と 3 つのモデルとのフィッティングを行い気導パワーエンベロープと骨導パワーエンベロープ間の $MTF e_h^2(t)$ モデルを以下の式で定義した

$$e_h^2(t) = a^2 \exp(-2bt). \quad (1)$$

このモデルはゲインを制御するパラメータ a と, 減衰特性を制御するパラメータ b を含んでいる. 先行研究で, パラメータ b の推定法は平松と鷗木によって提案されている [8]. パラメータ a については, 気導/骨導データベースを用いて気導パワーエンベロープと骨導パワーエンベロープの変換関係を解析した結果, 回帰曲線 $a_n = -cn^{-1} + d$ で近似できることが示唆された. ここで, n はフィルタバンクのチャンネル番号である. また, パラメータ c と d は観測点にのみ依存する. これによりパラメータ a と b を気導音声の情報が必要とすることなく決定することが可能となり, MTF に基づくブラインド骨導音声回復法が実現した.

最後に, シミュレーションにより提案法の評価を行った. 評価項目として, 音質評価尺度として, 対数スペクトル歪 (LSD), 線形予測係数を用いた LSD, ケプストラム距離, メルケプストラム距離. 音声明瞭度の評価尺度として音声明瞭度を考慮した LSD [9] を用いた. 評価の結果, 提案法の有用性が示された.

参考文献

- [1] 北森 進, 滝沢 正浩, “明瞭度試験による骨導音声の分析,” 電子情報通信学会論文誌, vol. J72-A, no. 11, pp. 1764–1771, Nov. 1989.

- [2] 加村 健一郎, 齊藤 裕, 福島 学, 石光 俊介, 柳川 博文, “振動ピックアップ型マイクによる収録音声の特性補正について,” 日本音響学会講演論文集, 1-Q-13, pp. 661-662, March 2002.
- [3] Shunsuke Ishimitsu, Hironori Kitakaze, Yasuyuki Tsuchibushi, Hirofumi Yanagawa, and Manabu Fukushima, “A noise-robust speech recognition system making use of body-conducted signals,” *Acoustical Science and Technology*, vol. 25, no. 2, pp. 166–169, 2004.
- [4] Tosiki Tamiya. and Tetsuya Shimamura, “Reconstruct Filter Design for Bone-Conducted Speech,” *Proc. ICSLP2004*, vol. II, pp. 1085–1088, October 2004.
- [5] 石光 俊介, 高家 陽介, 堀畑 聡, 北風 裕教, 柳川 博文, “適応フィルタを用いた骨導音明瞭度向上の基礎研究,” 日本音響学会講演論文集, 1-Q-23, pp. 681-682, March 2002.
- [6] Tat Thang Vu, Germine Seide, Masashi Unoki, and Masato Akagi, “Method of LP-based blind restoration for improving intelligibility of bone-conducted speech,” *Proc. Interspeech2007*, pp. 966–969, Antwerp, Belgium, August 2007.
- [7] Kenji Kimura, Masashi Unoki, and Masato Akagi, “A study on a bone-conducted speech restoration method with the modulation filterbank,” *NCSP05*, pp. 411-414, Honolulu, USA, March 2005.
- [8] Sota Hiramatsu, and Masashi Unoki. “A Study on the Blind Estimation of Reverberation Time in Room Acoustics,” *J. Signal Processing*, v10. 12, No.4 , pp. 323-326. July 2008.
- [9] ANSI S3.5-1997, “American National Standard Methods for Calculation of the Speech Intelligibility Index,” 1997.