

G-6

骨導マイクロホン及び咽喉マイクロホンをを用いた音声認識に関する研究 - 骨導マイクロホン及び咽喉マイクロホンの周波数特性 -

A study on speech recognition using a bone-conduction microphone and a throat microphone - Frequency response of a bone-conduction microphone and a throat microphone -

○藤岡紘展¹, 関弘翔², 細野裕行²*Hironobu Fujioka¹, Hiroto Seki², Hiroyuki Hosono²

Abstract: In this study, we aim to improve the performance of speech recognition using a bone-conduction microphone and a throat microphone. In this report, we measured the frequency response of the bone-conduction microphone and the throat microphone using the frequency sweeping sound and the white noise. In addition, we compared their frequency responses and the frequency response of the air-conduction microphone.

1. まえがき

近年、交通・工事騒音などの増加に伴い、高騒音環境下における音声伝送の手段として、環境雑音の低減性に優れた骨導マイクロホン及び咽喉マイクロホンの利用が提案されてきた^[1]。また、音声認識デバイスの普及から、骨導マイクロホンなどを利用した音声認識の需要が今後見込まれる。しかし、骨導マイクロホンなどは、気導マイクロホンに比べて収音できる帯域が狭く特性も異なるため、気導マイクロホンを想定した通常の音響モデルでは音響ミスマッチによる音声認識の性能低下が生じる^[2]。

そこで本研究では、骨導マイクロホン、咽喉マイクロホンを使用時の音声認識の性能向上を目的とする。本報告では、周波数をスイープした音源とホワイトノイズの音源を入力信号とし、骨導マイクロホン及び咽喉マイクロホンの周波数特性を測定し、気導マイクロホンの周波数特性と比較検討した。

2. 実験方法

骨導マイクロホン及び咽喉マイクロホンの周波数特性を明らかにするために周波数解析(DFT)を行う。ここではその為の実験方法を示す。

本実験で使用するサンプル音源の仕様を以下に示す。また、使用機材を Table 1.に、システム図を Fig. 1 に示す。

[サンプル音源]

- ・サンプリング周波数 : 192kHz
- ・ビット数 : 24bit
- ・ファイル方式 : wav
- ・周波数スイープ音(60 秒で 50Hz~3kHz まで)
- ・ホワイトノイズ(60 秒)

本実験では、音声の周波数の範囲が 6kHz まで^[3]、人間の可聴域が 20Hz~20kHz、気導マイクロホンの周波数特性が 50Hz~15kHz^[4]、骨導マイクロホン及び咽喉マイクロホンの周波数特性が 200Hz~3kHz^{[6][7]}であることから、上記の仕様のサンプル音源を入力信号としてそれぞれのマイクロホンで録音した音について周波数解析を行う。解析対象音は上記のサンプル音源を入力信号として、気導スピーカーで流した音を気導マイクロホン(Table 1(a))で録音した音(以下、気導音とする)、骨導スピーカー(Table 1(b))で流した音を骨導マイクロホン(Table 1(c))を装着して録音した音(以下、骨導音とする)、同じく骨導スピーカーで流した音を咽喉マイクロホン(Table 1(d))を装着して録音した音(以下、咽喉音とする)の3つとする。

Table 1. Used equipment

	
(a) Air-conduction microphone (I456 Microphone) ^[4]	(b) Bone conduction Speaker (WR-3 CL-1001) ^[5]
	
(c) Bone-conduction microphone (EN21N-Tip) ^[6]	(d) Throat microphone (SH-12jK) ^[7]

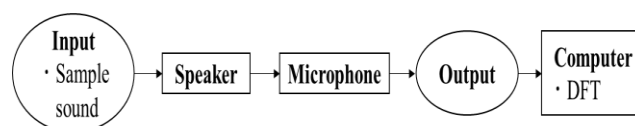


Figure 1. System diagram

3. 結果

ここでは、気導・骨導・咽喉マイクロホン各々の周波数解析の結果を述べる。

3.1. 周波数スイープ音源

周波数スイープ音源での気導音、骨導音、及び咽喉音における周波数解析の結果を Figs. 2~4 に示す。なお、波形は各周波数における音圧の最大値を示す。

Fig. 2 から気導マイクロホンの周波数特性は±2dB のフラットな特性であることがわかる。Fig 3 から骨導マイクロホンの周波数特性は約 400Hz 付近でピークを持ち、2000Hz 以降で徐々に減衰することがわかる。また、Fig 4 から咽喉マイクロホンの周波数特性は 200Hz、1550Hz 付近でピークを持つことがわかる。ここで、200Hz 付近でピークを持つのは、人の呼吸音(約 180Hz) であると考えられる^[8]。

3.2. ホワイトノイズ

ホワイトノイズでの気導音、骨導音、及び咽喉音における周波数解析の結果を Figs. 5~7 に示す。なお、波形は各周波数における音圧の最大値を示す。

Fig. 5 から気導マイクロホンの周波数特性は周波数スイープ音源と同様にフラットであることがわかる。Fig. 6 から骨導マイクロホンの周波数特性も同様に 400Hz 付近でピークを持ち、2000Hz 以降で徐々に減衰することがわかる。また、Fig. 7 から咽喉マイクロホンの周波数特性も同様に 200Hz、1550Hz 付近でピークを持つことがわかる。

4. まとめ

本報告では、骨導マイクロホン、咽喉マイクロホンの周波数特性を検討した。気導マイクロホンの周波数特性に比べ、骨導マイクロホンの周波数特性は 400Hz 付近でピークを持つこと、2000Hz 以降で徐々に減衰することが明らかになった。また咽喉マイクロホンの周波数特性は 200Hz、1550Hz 付近でピークを持つことが明らかになった。

今後は骨導マイクロホンや咽喉マイクロホンを使用した音声認識について検討し、認識性能の向上を目指す。

5. 参考文献

- [1]. 前田秀彦, 他:「骨固定型ピックアップから導出した直接骨導音の音響特性」, 音声医学言語学会, 2016
- [2]. 林升柯, 他:「複数の装着型マイクを用いた多人数会話音声認識に関する検討」, 情報処理学会, 2016
- [3]. 城戸健一, 「音響工学」, コロナ社
- [4]. Mic W ホームページ, <http://www.mic-w.com/product.php?id=24> (2018-09)
- [5]. earsopen ホームページ, https://earsopen.jp/product/wr_3_cl1001/ (2018-09)
- [6]. TEMCO ホームページ, http://www.temco-j.co.jp/pr_hg42tb/ (2018-09)
- [7]. 喉頭マイク (咽喉マイク), https://www.nanzu.jp/syohin/keitai_year.htm (2018-09)
- [8]. 宮西寛奈:「ウェーブレット変換を用いた呼吸異常音の特徴抽出」, 高知工科大学学士学位論文, 2014

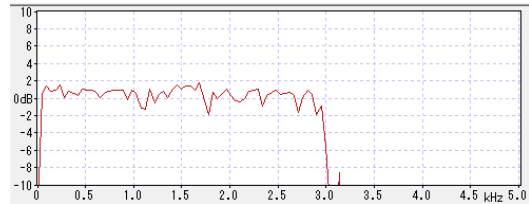


Figure 2. Sweep of air conduction sound

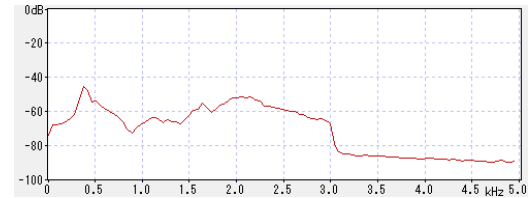


Figure 3. Sweep of bone conduction sound



Figure 4. Sweep of throat conduction sound

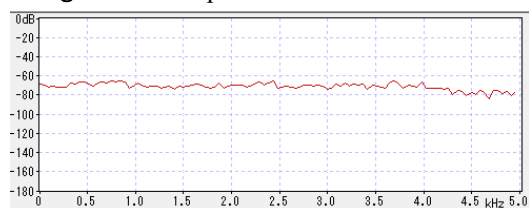


Figure 5. White noise of air conduction sound

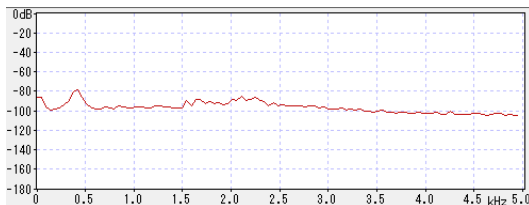


Figure 6. White noise of bone conduction sound

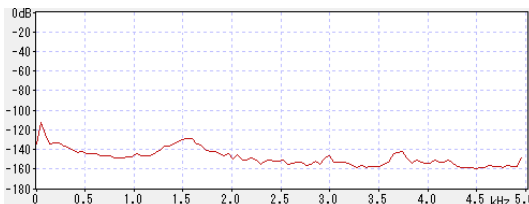


Figure 7. White noise of throat conduction sound