

[研究助成 (B)]

骨導デバイスを利用した聴覚拡張現実の実現

2241901

研究代表者
(助成金受領者)北陸先端科学技術大学院大学
先端科学技術研究科

教授

鷗 木 祐 史

共同研究者

北陸先端科学技術大学院大学
先端科学技術研究科

特任准教授

木 谷 俊 介

北陸先端科学技術大学院大学
先端科学技術研究科

特任助教

上江洲 安 史

[研究の目的]

拡張現実とは、現実世界とデジタル情報を融合させることにより、快適で豊かな社会を実現する技術として期待されている。また、人間拡張は、人間の能力を補完・向上させ、あるいは新たに獲得するための技術であり、「人機一体」ともいえるように、技術と人間の能力が自然に融合して一体化することが求められる。そのため、これらの技術はまさに「人間と機械の調和」が求められる重要なものである。

音を利用した聴覚拡張は、我々の聴き取りをサポートし、さらには我々の聴取能力を補完・向上させ、安心安全な生活を快適に送ることができる。その際、音の提示方法を注意深く検討しなければならない。通常はヘッドホン提示が考えられるが、耳を閉塞するため外界の音を遮断してしまうことになる。耳を開放させた状態で音を提示できるものとして骨導デバイスがある。これを使った場合、現実世界の音を気導音で、デジタル情報を骨導音で提供することにより、聴覚で融合させることができ、拡張現実を実現できる。

一方、聴覚拡張では、聴取能力の補償として補聴器のような聴き取りの改善が求められる。骨導提示は、伝音性難聴に対しては大きな聴き取り能力の改善効果を生むことが知られているが、骨導音声の音質や明瞭性・了解性を低下さ

せてしまう。また、気導音と骨導音の併用は、その知覚的融合の効果から音の強調や抑圧を可能とし、さらには聴覚的注意を促すことが期待される。そのため、骨導デバイスを利用した音の提示方法は、拡張現実だけでなく聴覚拡張を実現するための重要な要素技術になる。

本研究では、聴知覚メカニズムに着目し、骨導提示による音の強調と気導・骨導音の融合による音のエンリッチメント（聴覚拡張）を検討した。さらに、骨伝導を触覚知覚ととらえた聴覚代行の実現可能性を検討した。本稿では次の3つの課題について報告する。

1. 骨導提示音の音質・了解度改善法の構築
2. 骨導提示による気導音の強調・抑圧法の構築
3. 触覚による聴覚代行の実現可能性

[研究の内容、成果]

1. 骨導提示音の音質・了解度改善法の構築

本研究では、図1に示すように、骨導提示音声の了解度改善法を提案した。まず、一次高域強調処理(RT-FOE, 図1(a))は、気導音声を基準とした側頭部振動の伝達特性から逆特性を用いて、一次近似したFIRフィルタを利用することで、骨導伝達による直線的な高域減衰を補償するものである。

次に、子音強調処理(CE-IMP, 図1(b))は、骨導音声において子音の知覚の難易性が高くな

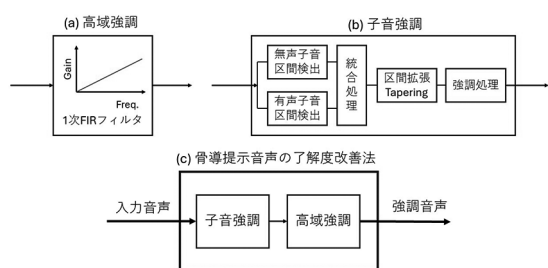


図1 骨導提示音声の了解度改善法

ることから、子音部分を強調することで難易性を下げることが狙いとしたものである。ここでは、無声子音・有声子音のそれぞれで子音区間検出を行い、区間の統合・後処理を施した上で、子音部分を強調することで実現した。

最後に、両処理を組み合わせる方法（RT-FOE+CE-IMP, 図1(c))を検討する。子音強調処理における無声子音・有声子音区間検出では閾値処理を採用しており、一次高域強調処理がその閾値処理に影響する可能性がある。そのため、本稿では、子音強調処理、一次高域強調処理の順に実行することにした。

三つの処理の効果を評価するために、骨導提示音声の了解度試験を行った。実験には、正常聴力を有する20代の日本語母語話者20名（男性10名、女性10名）が参加した。骨導提示音声として、親密度別単語了解度試験用音声データセット（FW07）の音声を使用した。各単語は4モーラで構成され、それぞれに4段階の親密度が設定された。本実験ではすべての親密度ランクの単語が含まれるようにした。

音声刺激として、3種類の強調処理（RT-FOE, CE-IMP, RT-FOE+CE-IMP）を適用した音声と、強調処理なし（No emphasis）= 原音声を用了。雑音環境を模擬するため、2種類の提示音圧レベル（55 dBと75 dB）のピンク雑音を用了。条件の総数は32（強調タイプ4種類×親密度4ランク×雑音レベル2種類）であった。

実験参加者には、ラウドスピーカからピンク雑音を気導提示すると同時に、骨導デバイスから音声刺激を骨導提示し、聴き取った単語を

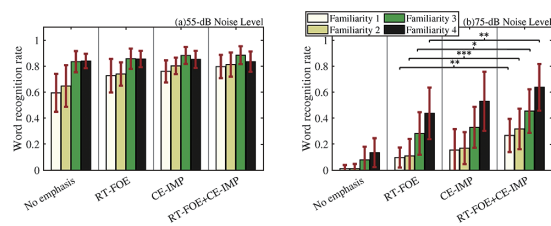


図2 強調タイプ・親密度別の単語了解度

キーボードでカタカナ入力させた。設問は1条件あたり20単語、計32条件640単語とした。全参加者が同一の単語を聴取することを避けるために、含まれる単語が互いに異なるように設計された単語リストを8セット用意した。ある単語について、4モーラすべてを正しく入力した場合にのみ、その単語を正答とした。親密度ランク、強調タイプごとの単語正答率を算出し、比較評価に用いた。

図2に、雑音レベル (a) 55 dB, (b) 75 dB の条件下における骨導提示音声の単語正答率の平均および標準誤差を示す。単語正答率に対する三要因分散分析の結果、強調タイプ [$F(3, 57) = 200.9, p < .001$] の主効果が認められた。さらに、三要因の交互作用が有意 [$F(9, 171) = 9.72, p < .001$] であったことから、雑音レベルごとに結果を分けて検討した。

雑音レベル75 dBでの結果について、多重比較検定を行った。検定結果はHolm法で補正された。RT-FOE+CE-IMPは強調処理なしおよびRT-FOEと比べて、すべての親密度ランクにおいて有意に正答率が高いことを示した。一方で、RT-FOE+CE-IMPはCE-IMPと比べて全体的に高い正答率の傾向を示したが、統計的に有意な差は認められなかった。

RT-FOE+CE-IMPは、雑音レベル75 dB条件下のすべての親密度ランクに対して、RT-FOEよりも有意に高い単語正答率を示した。この結果は、骨導提示で減衰しやすい高周波数成分の補償に加えて、子音部の頑健な検出・強調処理を施すことで、子音の識別手掛かりが増強され、高騒音環境下でも高い了解度をもたら

したことを意味している。

一方、CE-IMP との比較では、いずれの親密度ランクにおいても高い単語正答率を示す傾向があるが、その差は統計的には有意ではなかった。このことから、CE-IMP 単独の処理でも、単語の識別手掛かりがすでに効果的に補償されていた可能性がある。

今後は、異聴分析や音節ごとの評価を実施し、提案法の効果と限界点を詳細に検討する予定である。

2. 骨導提示による気導音の強調・抑圧法の構築

騒音抑圧法として、耳栓やイヤーマフといった聴覚保護具があるほか、特定音の騒音の抑圧技術としてアクティブノイズコントロールがある。しかし、これらの方法はいずれも耳を塞ぐことを想定しているため、雑音だけでなく、聞き逃してはいけない音声まで抑圧してしまう。そのため、耳を塞がずに外界の騒音を抑圧する方法の実現が期待されている。

本課題では、骨導提示音の振幅・位相を操作することで気導音を知覚的に抑圧できるか検討した。

この方法の特徴は、耳を開放した情報で、特定の音を抑圧できることである。ここでは、7 名の実験参加者に対して A 特性音圧レベル 50 dB、周波数が 250 Hz、500 Hz、1 kHz、2 kHz、4 kHz、8 kHz の純音を利用して、気導・骨導音のラウドネスが等価になるように振幅を決定し、気導・骨導音を同時提示したときに最もラウドネスが小さくなるような位相を決定した。

提案法を利用して、気導音の抑圧効果を調べたところ、図 3 のようになった。7 名中 5 名で抑圧の効果がみられ、2 名の参加者で若干の増幅の効果がみられた。また、4 kHz の純音では全参加者で抑圧の効果がみられた。これらの結果から、骨導音による気導音の抑圧が可能であることがわかった。

今後は、対象を純音から複合音へ拡張して検討する。また、位相を逆相で処理することで増

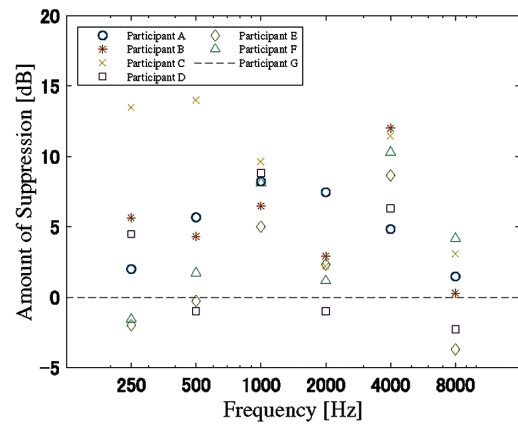


図3 骨導音による気導音の抑圧特性

幅効果も期待できるため、その特性も検討する予定である。

3. 触覚による聴覚代行の実現可能性

聴覚障がいとは、聴覚系の周波数選択性の低下や時間分解能の低下から音声コミュニケーションの困難さを招き、結果としてQoLを低下させてしまう。この課題を解決するためには、聴覚機能を補完する技術や他の感覚器によって聴覚機能を代行する技術の開発が必要である。代表的な聴覚代行として、触覚を用いた聴覚代行がある。これは、触覚の受容器が身体表面全域に分布し、聴覚と同様に、振動から情報を取得可能なためである。例えば、タクトイルボコーダ（音声のフォルマント情報を空間的パターンと振動の強弱として触覚提示するもの）がある。これは言語情報知覚の可能性を示したが、その仕組み上、感情といった非言語情報を知覚させることは困難である。

従来、触覚知覚は、聴覚と比べて非常に低い周波数をもつ振動を知覚するものである。しかし、高周波振動であっても、触覚の変調知覚として、振動強度の時間変動パターンを知覚していることも知られている。一方、我々の研究グループにより、聴覚の変調知覚の検討から、音声の時間振幅包絡線情報に非言語情報が含まれていることが知られている。もし同じアナロジーとして、触覚の変調知覚として音声刺激の

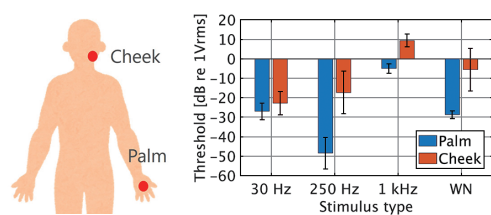


図4 掌と頬の触覚閾値

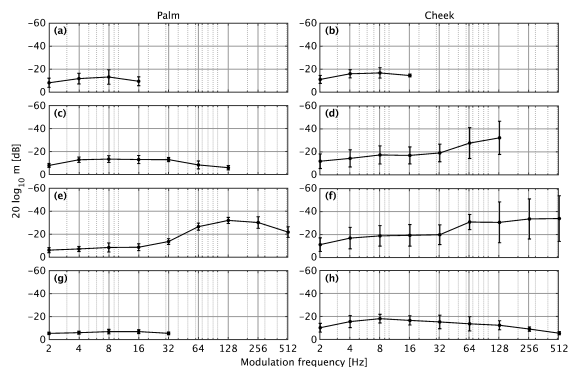


図5 触覚知覚における時間変調伝達関数。左は掌、右は頬での結果、パネルの上から下に向かって周波数 30 Hz, 250 Hz, 1 kHz, 白色雑音での結果

振幅包絡線情報を検知できているとしたら、触覚の変調知覚として非言語情報を知覚できるかもしれない。

本課題では、このような仮説のもと、音声に触覚刺激として提示した際の非言語情報知覚の実現可能性について検討する。

まず、触覚知覚の検知閾を調査した。実験には正常聴力を有する日本語母語話者5名（大学院生23～26歳）が参加した。男性3名（右利き1名、左利き2名）、女性2名（右利き2名）が参加した。

図4に、各提示部位および刺激条件における触覚閾値を示す。閾値は、基準電圧1 Vrmsに対する出力電圧をdB換算した値である。掌での触覚閾値は、250 Hzで最も低く、1 kHzで最も高かった。これに対し、頬での触覚閾値は、30 Hzで最も低く、1 kHzで最も高かった。

次に、同じ実験参加者に対して、触覚の変調知覚として時間変調伝達関数（TMTF）を調査した。図5にその結果を示す。TMTFが一種の帯域通過特性のように、ある特定の変調周

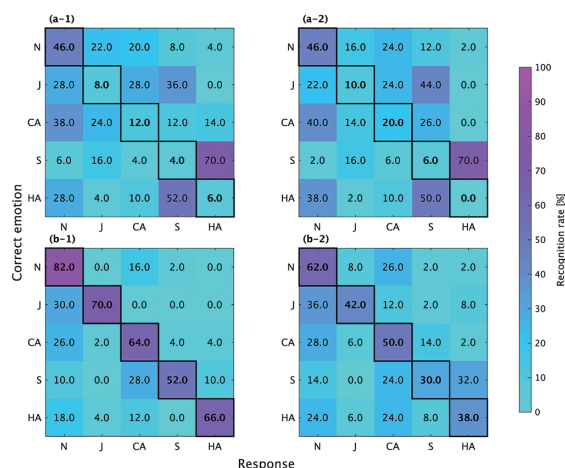


図6 感情知覚実験の結果：提示部位が（a）掌と（b）頬、刺激が（1）原音声と（2）雑音駆動音声の場合

波数帯域に感度をもつことがわかる。特に図4（a-3）は、Lamoreらの報告と整合する結果である。これらの結果から、触覚の変調知覚は、概ね信号の変調成分（32 Hz以下）を検知するような特性であることが予想される。

最後に、同じ実験参加者に対し、触覚刺激提示による音声の感情知覚実験を行った。図6に結果を示す。図は回答の混同行列であり、横は提示した触覚刺激の感情ラベル、縦は実験参加者の回答に対応する。図中のN, J, CA, S, HAはNeutral, Joy, Cold anger, Sadness, Hot angerを表す。対角成分は感情認識率、非対角成分は誤認識を表す。頬への刺激提示では、原音声および雑音駆動音声のいずれの場合も、全ての感情で認識率がチャンスレベルを上回った。頬は掌に比べて高い感情認識率を示す傾向があった。同部位において原音声とNVSで類似した回答傾向がみられたことから、触覚による感情知覚の手掛かりは両刺激で共通していた可能性がある。

〔成果の発表、論文など〕

- [1] 鶴木祐史, 進化するヒトと機械の音声コミュニケーション vol.2 AIの活用と感情に寄り添う音声認識・合成の新展開, “骨導デバイスを利用した音声コミュニケーション: 人と機械による音声認識” (第1編, 第2章, 第2節), エヌ・ティー・エス, 2025.

- [2] Imam Nurul Hukmi, Quoc-Huy Nguyen, **Masashi Unoki**, "Quantifying the Contribution of WORLD Vocoder Features to Bone-Conducted Speech Enhancement," Proc. 2026 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing, pp. 53-56, March 2026.
- [2] 牧野哲平, 上江洲安史, 木谷俊介, **鵜木祐史**, "触覚刺激提示による音声の非言語情報知覚の検討," 日本音響学会 2026 年春季研究発表会講演論文, 1-R-3, pp. 707-710, 日大理工, March 2026.
- [3] Quoc-Huy Nguyen, **Masashi Unoki**, "Low-Latency, Unwrap-Free, and Differentiable Phase Features for Speech Analysis," Proc. ASJ 2026 Autumn Meeting, 1-Q-39, pp. 1095-1098, Nihon University, March 2026.
- [4] 中川怜奈, 磯山拓都, 上江洲安史, **鵜木祐史**, "一次高域・子音強調処理による骨導提示音声の了解度改善," 日本音響学会 2026 年春季研究発表会講演論文, 3-P-15, pp. 669-670, 日大理工, March 2026.
- [5] 牧野哲平, 上江洲安史, 木谷俊介, **鵜木祐史**, "触覚の知覚特性に基づく非言語情報知覚の検討," 日本音響学会聴覚研究会資料, Vol. 56, No. 2, H-2026-53, pp. 97-102, Feb. 2026.
- [6] 牧野哲平, 上江洲安史, 木谷俊介, **鵜木祐史**, "触覚刺激提示による感情音声認識の検討," 日本音響学会聴覚研究会資料, Vol. 55, No. 6, H-2025-77, pp. 415-420, Nov. 2025.
- [7] 牧野哲平, 上江洲安史, 木谷俊介, **鵜木祐史**, "触覚刺激提示による感情音声認識の予備的検討," 2025 年度電気・情報関係学会北陸支部連合大会 H3-1, 富山大学, Sept, 2025. 【日本音響学会北陸支部優秀学生賞受賞】
- [8] 幸田怜奈, **鵜木祐史**, 小澤賢司, 鳥谷輝樹, "骨導提示音による気導提示音抑圧のための振幅・位相条件の検討," 情報処理学会第 87 回全国大会発表資料, 7T-08, pp. 2-621-2622, 立命館大学, March 2025.
- [9] 幸田怜奈, **鵜木祐史**, 小澤賢司, 鳥谷輝樹, "開放耳における気導提示純音の骨導提示音による抑圧効果," 日本音響学会 2024 年秋季研究発表会講演論文, 3-P-37, pp. 849-850, 関西大学, Sept. 2024.