

3-6-1

3-6-1 吃音者音声合成のための連発・難発を考慮した合成モデル学習

Training of speech synthesis model considering repetitions and blocks for stutterer-oriented text-to-speech

©高島遼一^{1,2}, △安井美鈴³, 滝口哲也¹
(¹神戸大学, ²JST さきがけ, ³大阪人間科学大学)

- ◆本研究では吃音当事者の声質でかつ流暢な音声を合成可能なテキスト音声合成システムを検討する。
- ◆吃音音声を用いて合成モデルを学習するとき、台本のテキストと実際の吃音イベントが含まれた発話との間で内容のミスマッチが生じるため、生成される音声も非流暢なものとなる。
- ◆本発表では、吃音音声を手動で書き起こすことなく、台本のテキストを自動で修正しながら合成モデルを学習する手法を提案する (Fig. 1)。
- ◆実験の結果、提案手法はベースの音声合成モデルの合成品質を劣化させずに、合成音声の流暢さを改善できることが示された。

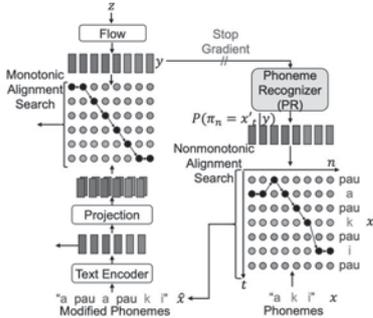


Fig.1: The proposed model training procedure.

3-6-2

3-6-2 言語非依存な口真似データセット構築と口真似のみからのTransformerによる効果音合成

Construct a language-independent voice mimicry dataset and Synthesis sound effects only from the voice mimicry with Transformer

☆滝沢力, △平井重行(京産大)

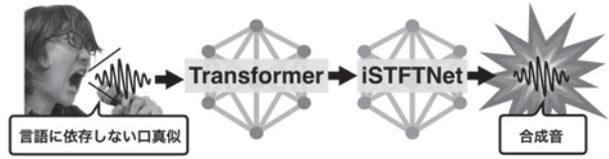


Fig.1: 口真似のニュアンスを反映する音響合成手法

- ◆目的: 言語非依存な口真似の特徴を反映させた効果音合成技術の確立
- ◆提案法: 音そのものを真似た口頭表現の使用 (Fig.1 参照)
 - 口真似から効果音への変換を、それらのメルスペクトログラムを入力として Transformer で学習
 - iSTFTNet でメルスペクトログラムから波形合成
- ◆データセット: 爆発音にフォーカスして構築
 - 1,748 種類の爆発音に対し、3名で7,768 回口真似実施
- ◆モデルの学習
 - 構築データセットで Transformer, iSTFTNet を学習
- ◆合成の検証
 - 口真似のニュアンスが反映された合成結果を確認

合成音聞けます! →



3-6-3

3-6-3 J-CHAT: 音声言語モデルのための大規模日本語対話音声コーパス

Japanese Large Scale Dialogue Corpus for Human-AI Talks

©中田亘, 関健太郎, 谷中瞳, 齋藤佑樹(東大), 高道慎之介(慶大/東大), 猿渡洋(東大)

1. 風吹(1)名無し: 2024/09/06(金) 12:01:23.45 ID:b0d1234

AI と人間の会話って、最近すごく話題になってるよなあ。特に音声の奴。実際どうなん?

2. 風吹(2)名無し: 2024/09/06(金) 12:02:10.77 ID:efgh5678

正直、ちょっと怖くね? そのうち全部 AI に話しかけるようになるんじゃない?

3. 風吹(3)名無し: (水) 12:02:45.12 ID:ijkl9012

>>2 もうすでにそうなってる感あるけどなwww

4. 風吹(4)名無し: 2024/09/06(金) 12:03:12.22 ID:mnop3456

ところでなんか「J-CHAT」ってやつ公開されたいいな

5. 風吹(5)名無し: 2024/09/06(金) 12:03:55.78 ID:qrst7890

>>4 それ何? 聞いたことねーわwww

6. 風吹(6)名無し: 2024/09/06(金) 12:04:30.67 ID:uvwx1234

>>5 どうやら7万時間日本語音声コーパスらしい

7. 風吹(7)名無し: 2024/09/06(金) 12:03:55.78 ID:qrst7890

>>6 クソでかじゃんワロタwww

3-6-4

3-6-4 UTMOSv2: 自然性 MOS 予測におけるスペクトログラム特徴量と SSL 特徴量の統合的利用

UTMOSv2: Integrating Spectrogram and SSL Features for Naturalness MOS Prediction

☆馬場 凱渡, 中田 亘, 齋藤 佑樹, 猿渡 洋(東大)

3-6-5

3-6-5 合成音声の大規模主観比較評価からの絶対品質推定

Absolute quality estimation of synthetic speech from large-scale subjective preference evaluations.

◎安田裕介, 戸田智基 (名古屋大学)

本研究は、合成音声の大規模主観比較評価から、統計的な手法で絶対品質値を推定します。これによって、従来の MOS 評価値に対して、バイアスの少ない絶対品質値を提案します。

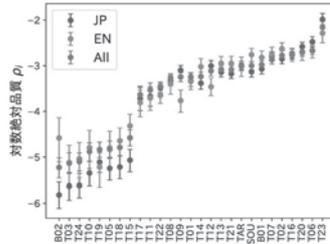


図1: 比較評価結果から推定した対数絶対品質

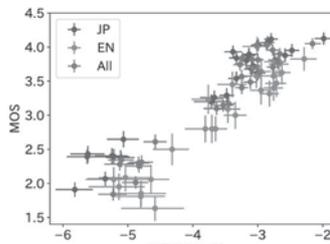


図2: 推定した対数絶対品質と MOS との比較

3-8-1

3-8-1 非接触音響探査法による耐火壁内の欠陥検出に関する検討

Study on defect detection in fireproof walls using noncontact acoustic inspection method

○杉本恒美, 上地樹, 杉本和子, 中川裕 (桐蔭横浜大院),
△山村一生, △原田朋弘 (三菱重工環境・化学エンジニアリング株),
△石岡昌人 (三菱重工業株)

- ◆ごみ焼却プラントを想定した耐火材模擬供試体および実炉壁に対して、非接触音響探査法を用いた耐火壁内の欠陥検出に関する実験的検討を行った。
- ◆供試体の振動エネルギー比(VER: Vibrational Energy Ratio)による映像結果例を Fig.1(a)に示す。叩き点検では検出できなかった深さ 36~70 mm の深さに埋設された 400 mm 角の模擬欠陥 (図中右上の白破線内) が検出できることが確認された。
- ◆実炉壁における振動エネルギー比による映像結果例を Fig.1(b)に示す。白破線部より上が叩き点検により欠陥と判定された領域であり、叩き点検とほぼ同等の結果が得られることが明らかになった。

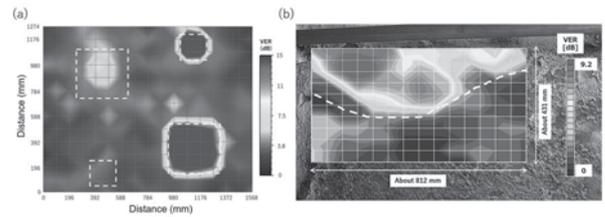


Fig.1: Example of experimental results based on vibrational energy ratio, (a) fireproofing specimen, (b) actual furnace wall.

3-8-2

3-8-2 XR 技術を用いた三次元音圧分布可視化手法に関する研究 -その1 定位精度向上を目的とした可視化表現の改良-

Study on visualization method of 3D sound pressure distribution using XR technology -Part 1: Improving visualization representation to enhance localization accuracy-

◎井上敦登, 寺岡航, 及川靖広(早大理工/INSPIREI)

佐藤考浩, 小林真人(飛鳥建設)

- ◆XR 技術と MVDR ビームフォーマにより、三次元の音圧レベル分布を実空間上で可視化する「音場可視化システム OTOMIRU™」を開発
- ◆計測毎のカラーマップ誤差が大きく、音場定位が困難な場合があった。
- ◆音源定位の容易化を目的に、計測データを設定した回数分過去まで移動平均化する設定や、計測間を時間補間する設定を追加。
- ◆無響室内スピーカの可視化実験にて、定常的な音は計測毎カラーマップの誤差が低減、突発的な音は Fig.1 のようにスピーカ位置の音圧が強いことを表すカラーマップを認識可能な時間表示することを実現

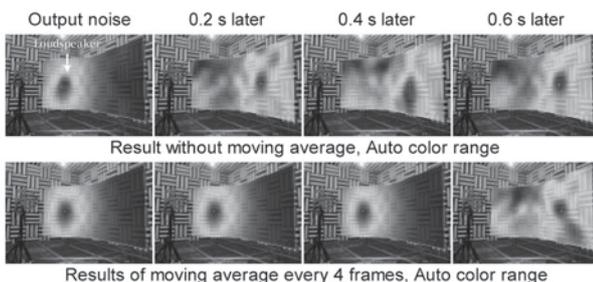


Fig.1 Visualization results of a loud-speaker with momentary white noise.

3-8-3

3-8-3 XR 技術を用いた三次元音圧分布可視化手法に関する研究 -その2 改良した可視化表現の精度検証-

Study on visualization method of 3D sound pressure distribution using XR technology -Part 2 Verification of the accuracy of the improved visualization representation-

◎佐藤考浩, 小林真人(飛鳥建設),

及川靖広, 井上敦登, 寺岡航(早大理工/INSPIREI)

- ◆本研究では、三次元の音圧レベル分布を XR 技術により実空間上で可視化する「音場可視化システム OTOMIRU™」(開発システム)を開発している。本稿では、音源定位精度の向上を目的として移動平均した音圧レベル分布を表示させるなどの改良を行った可視化表現の精度を検証するため、建物内部の音響性能を計測する音場を開発システムにより可視化した結果を報告する。
- ◆重量床衝撃音を可視化した結果、移動平均せずに表示させた結果では音圧レベル分布にばらつきがみられる一方、8 フレーム分の音圧レベル分布を移動平均した結果では、天井面や壁面で音圧レベル分布が大きくなっている箇所を目視で特定することができた (Fig.1)。

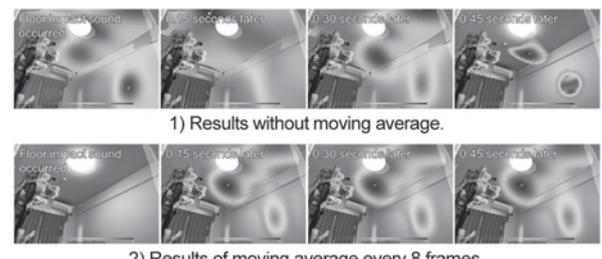


Fig.1 Visualization results of heavy-weight floor impact sound.

3-8-4

3-8-4 複数音場モデルの結合を用いた複合現実技術による広範囲音場の可視化

Mixed reality visualization of broad sound field using the concatenation of multiple modeled sound fields

©内田彩芽, 津國和泉, 池田雄介(東京電機大), 及川靖広(早大理工)

◆背景

- ・ これまでに室形状情報と音場モデルを用いた複合現実技術 (MR) による室内インパルス応答の可視化を提案
- ・ マイクロホンから離れるにつれてモデル化推定精度が低下するため、可視化範囲の拡大が困難

◆提案手法

- ・ スパース等価音源法と虚像法に基づき直接音と反射音をモデル化
- ・ 複数の局所音場を結合し、広範囲音場を再構築
- ・ MR デバイスを用いて広範囲音場を現実空間に可視化

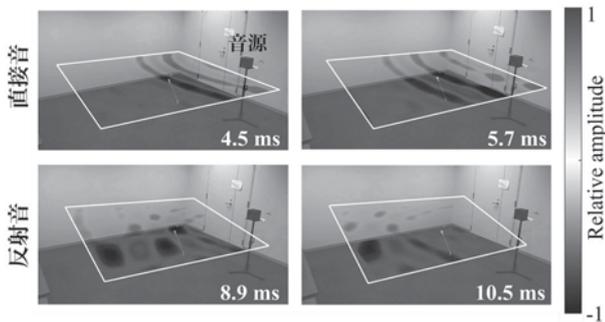


Fig.1: Visualization result of room impulse response using Mixed Reality

3-9-1

3-9-1 RC とせっこうボードの二重壁における遮音性能に関する実験 - 空気層を大きくとれる場合 -

Experimental study on sound insulation performance of double walls with RC and gypsum board - Case of large air layer -

©平岡千春, 吉岡清(佐藤工業)

- ◆RC 壁に対してせっこうボードなどの内装壁を付加する「ふかし壁」において、空気層の共鳴現象により低音域の遮音性能が低下することが知られている。通例、空気層の厚さは限られた範囲で計画・施工されるが、これを、大きくとれる場合、共鳴周波数が低域に移動することにより、低音域においても遮音性能を確保できる可能性がある。

- ◆空気層厚さが220mmと330mmの場合、せっこうボード及びグラスウール仕様の異なる試験体 (Fig. 1) に対して遮音実験を行い、その効果を検討した。その結果、低音域での遮音性能の低下は見られないことが確認された (Fig. 2)。

- ◆さらに、実験対象とした空気層厚さに対応する低域共鳴周波数が表れているかを、ハンマ加振による振動測定で確認したところ、計算値に近い周波数で応答が確認された。

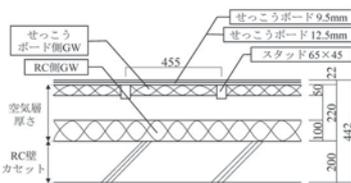


Fig.1 Example of test specimen

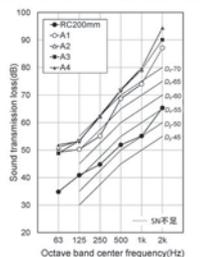


Fig.2 Experimental results of sound transmission loss

3-8-5

3-8-5 心疾患のある心拍音の可視化

Visualizing anomalousness of heartbeat from sound recorded through a stethoscope

☆松本真太郎(阪産大院), △塩安佳樹(はごろも内科・小児科), 高橋徹(阪産大院)

【目的】聴診器からの心拍情報を可視化し診察を支援

【手法】連続する5周期の周期同期相関値を求め 5×5の行列表現により周期性・非周期性をヒートマップで表現

【結果】疾患のない被験者1名と疾患のある被験者4名の心拍音波形を分析

- ・ 心疾患あり⇒非対角要素の値が大きい
- ・ 心疾患なし⇒非対角要素の値が小さい

心疾患の有無を視覚的に提示可能になった

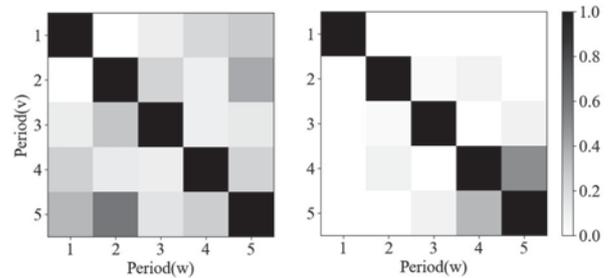


Fig.1: Heatmap of heartbeat (Left : without disease , Right : with disease).

3-9-2

3-9-2 界壁における木質系下地の仕様が遮音性能へ及ぼす影響

The influence of specifications of wooden substrates in separation walls on sound insulation performance

©兵藤伸也, 山下祐, 小林真人(飛鳥建設), 安田洋介(神奈川大学)

- ◆界壁における木下地の仕様が音響透過損失に及ぼす影響を把握するため、下地構成の異なる試験体を用意し、音響透過損失、インピーダンスおよび音響加振時の振動加速度分布を測定した (Fig.1)。

- ◆音響透過損失のディップが試験体1では250 Hz帯域、試験体2と3では160 Hz帯域で確認され、これらの要因を検討した。

- ◆面材におけるインピーダンスの結果から、160 Hzから250 Hzにおいて音響透過損失の低下に対応する固有振動数があることを確認した。

- ◆振動加速度分布の結果から、木質二重壁の音響透過損失は柱の位置により決まる固有振動数で低下し、間柱や垂木の拘束による影響は小さいことが分かった (Fig.2)。

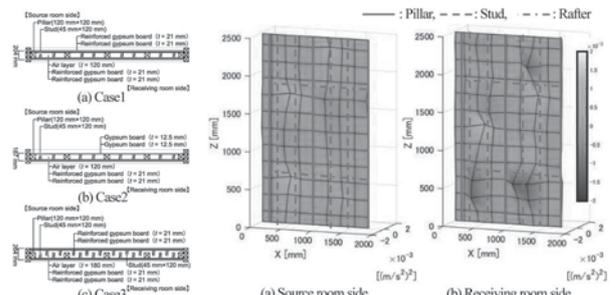


Fig.1 Specifications of experiment cases.

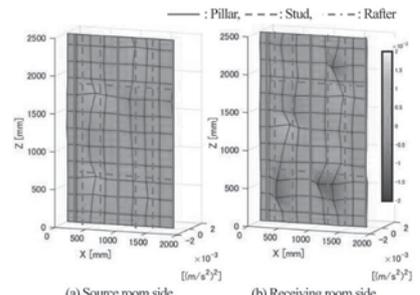


Fig.2 Example of vibration acceleration distribution at 249 Hz in case1.

3-9-3

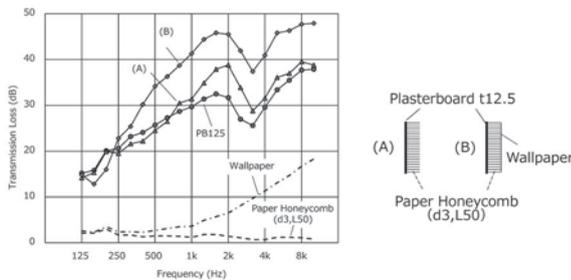
3-9-3 音波の入射角度制御による遮音性能向上手法に関する基礎研究

Basic study on methods for improving sound insulation performance by controlling the incident angle of sound waves

◎阿部将幸, 原澤悠, 小谷朋央貴, △江川隼太(フジタ)

部材表面にハニカム構造を設け、音波の入射角度を垂直方向に制御することによる遮音性能向上効果について、基礎的な検討を行い、以下の知見を得た。

- ◆部材表面にハニカム材を取り付けることにより、コインシデンス周波数帯域を中心に遮音性能が改善し、直径が小さいほど、長さが長いほど改善効果は増大する。
- ◆ハニカム材は部材に密着させて設置する方が効果的であり、10 mm以上のクリアランスが生じると遮音性能の改善効果が低下する。
- ◆ハニカム材は音波の入射側、放射側どちらに設置しても効果が得られる。
- ◆ハニカムの表面、及び内部に壁紙層を設けることで遮音性能の改善量が大きく向上し、周波数帯域も拡大する。



3-9-5

3-9-5 外周壁遮音性能の低周波測定法に関する数値解析的検討-室寸法の適用範囲の検証-

Numerical validation of the low-frequency procedure for façade sound insulation measurement-Verification of the applicability of room dimension-

◎劉金雨 (神奈川大・建築), 井上尚久(九大・芸工), 佐久間哲哉(東大・工)

- ◆木造試験家屋 (室容積 20 m³) における室内計測を行い、ISO 16283 低周波測定法の妥当性の検討、改良提案を行ってきたが、本研究では室寸法を変えた簡易部屋モデルを作成し、有限要素法による数値解析を行い、改良提案の適用性の検討を行った。
- ◆各部の音圧レベルについて、基本条件の数値解析結果は、木造試験家屋の実験結果と比較して、両方が傾向的に概ね対応することが見られた。数値解析モデルを用いてコーナー法の検討の有効性が確認できる。
- ◆奥行き方向の寸法を変えた Case 1a, 2a, 3a の3条件におけるコーナー法の検証結果をみると、ISO が規定されている室容積 25 m³ 未満の場合、80 Hz 帯域以下にコーナー法を用いた補正が必要である。一方、Case 3a (12 畳) の場合、80 Hz 以下では L_{2,all} に対する -1~2 dB の小差であるため、従来法を適用する可能性が示唆される。

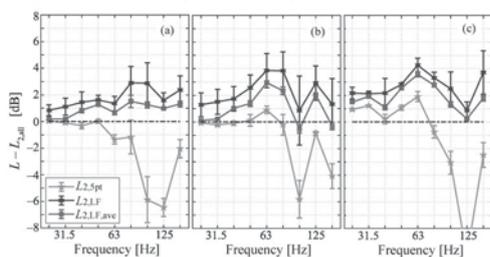


Fig.1: Average SPL by the ISO and corrected with average level of all corners: (a) 3.6×2.7×2.25; (b) 3.6×3.6×2.25; (c) 3.6×5.4×2.25.

3-9-4

3-9-4 多数の共鳴器を接続した内貼吸音ダクトの音響透過損失

Sound transmission loss of lined duct with multiple resonators

○轟羽琢元(竹中技研)

- ◆ダクト内を伝搬する音の減衰策として寸法の異なる多数の共鳴器をダクトに接続する方法が提案されている。本研究では、ダクトに接続された共鳴器への吸音材の挿入がダクトの音響透過損失に与える影響を調査した。

◆Fig. 1に本研究の対象である多数の共鳴器が接続されたダクトの断面図を示す。共鳴器の開口側および閉端側に吸音材を挿入した場合におけるダクトの音響透過損失の計算結果を Fig. 2 に示す。

- ◆共鳴器内に吸音材を挿入することで周波数特性がより平滑な透過損失が得られることが示された。

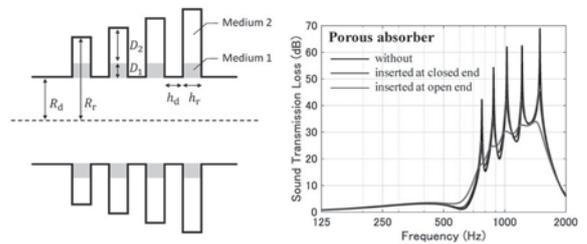


Fig.1: Cross section of circular duct with multiple resonators.

Fig.2: Effect of sound absorbers in resonators on sound transmission loss of duct with multiple resonators

3-9-6

3-9-6 三重周期極小曲面の遮音特性における入射角依存性に関する検討

The Dependence of Incidence Angle on the Sound Insulation Characteristics of Triply Periodic Minimal Surfaces

☆矢野拓実(近畿大院), 菅原彬子, 平栗靖浩, △堂寺知成(近畿大)

- ◆三重周期極小曲面 (TPMS) という、昆虫の鱗粉などで確認できる構造体は、特定の光に対してバンドギャップを生成する。これを音に応用することで遮音材として有用なのではと考えた。
- ◆有限要素法を用いて TPMS の遮音性能を検討する。既報[1]では垂直入射時の挿入損失を求めた結果、特定の周波数に対して高い遮音性能を示した。しかし、この遮音効果に入射角度依存性があるかを検討するために、斜入射時の挿入損失を検討する。
- ◆Fig. 1 に示す解析結果から、G 曲面の遮音効果には入射角依存性があることがわかり、P 曲面には、入射角依存のあるディップと、依存のないモードがあることがわかった。本報では、これらが起こる原因についても検討を行った。

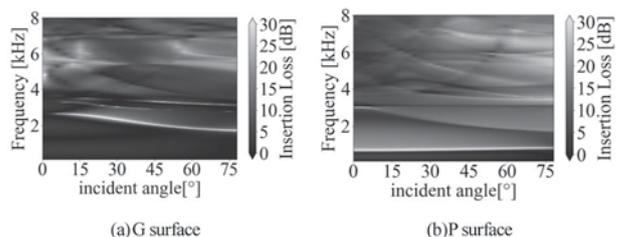


Fig. 1: Contour diagram when obliquely incident on the P and G surfaces

[1]矢野他, 音講論 (春), 465-468, 2024

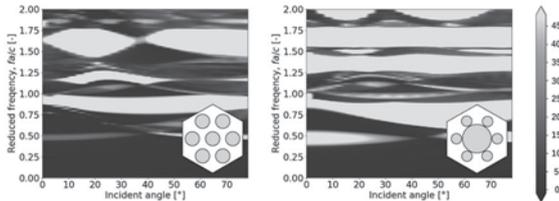
3-9-7

3-9-7 階層型ソニック結晶の遮音特性に関する基礎的検討

Sound Insulation Characteristics of Hierarchical Sonis Crystals

☆和泉誠大(近畿大院), 菅原彬子, 平栗靖浩(近畿大)

- ◆背景: ソニック結晶(SCs)は、周期構造により特定の周波数帯域に対してフォニックバンドギャップ(PBG)を形成する。PBGには入射波に対し角度依存性を持つこと、対象の周波数帯域が限定的であることなどといった課題がある。
- ◆目的: SCsの単位格子にフラクタル分割を用いることで階層型ソニック結晶(HSCs)の単位格子を作成する(Fig.1(a))。また HSCs の単位格子に対し形状の変化を加える(Fig.(b))ことで、単位格子の幾何形状と形成される PBG の関係性について検討を行う。
- ◆内容: (a)では規格化周波数 $f = 0.9$ にて角度依存性が小さくなることが確認された。また(b)との比較より、単位格子の形状に幾何学的変化を加えることで、挿入損失量が増加することと等方的 PBG が形成されることがわかった。本報では、充填率を揃えた SCs や異なる変化を加えたモデルとの比較検討を行い、遮音特性について分析した。



(a) Nomal HSCs (b) Changed HSCs
Fig.1 Contour map of insertion loss at oblique incidence

3-9-9

3-9-9 バフル板の振動を考慮した周期配置スリット共鳴器の吸音率予測における有限要素モデルの比較

Comparison of Finite Element Models for Predicting Sound Absorption Coefficients of Infinitely Periodical Slit Resonators with Baffle Vibrations

◎米澤 美桜, 井上 尚久(九大・芸工)

バフル板の振動を考慮した周期配置スリット共鳴器の吸音率解析を取り扱う。共鳴器のエネルギー吸収に関して、線形化ナビエ-ストークス方程式に基づくモデル (VTA model)、粘性・熱伝導境界層の影響をヘルムホルツ方程式の境界条件として与えるモデル (BC model) の二つを音響振動連成系へ拡張したものを取り上げ、スリット共鳴器の吸音率を有限要素法によりそれぞれ算出した。解析結果から、遷移周波数以上の解析では、BC モデルで十分な精度を持つことが示された。

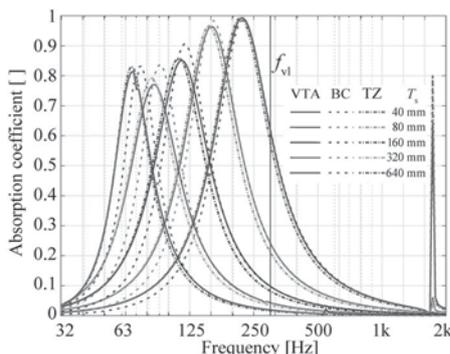


Fig.1: Normal incidence absorption coefficients calculated by using VTA and BC models for different combinations of spatial period T_s .

3-9-8

3-9-8 3次元ボロノイ分割を利用した拡散体の音響拡散特性に関する検討

The acoustic diffusion characteristics of a diffuser using three-dimensional Voronoi tessellation.

◎濱谷晃希(旭化成), 菅原彬子, 平栗靖浩(近畿大学)

- ◆背景: 周期を持たない形状を音響拡散体を利用することで、広帯域で無指向性に近い拡散特性が得られる可能性がある。
- ◆目的: ボロノイ分割を利用して非晶質な形状の音響拡散体を作成し、その音響拡散特性について検討する。
- ◆内容: 有限要素法を用いた数値解析より、散乱係数を算出する。3次元ボロノイ分割を利用した試料 (3D Voronoi) と、そこから斜面を取り除いた試料 (2D Voronoi)、規則的な試料 (Square grid) を作成し、比較検討を行う。
- ◆結果: 3次元ボロノイ分割を利用した試料 (3D Voronoi) は他の試料と比較して、広い周波数帯域で高い散乱係数を持つことが確認できた。(Fig.1)

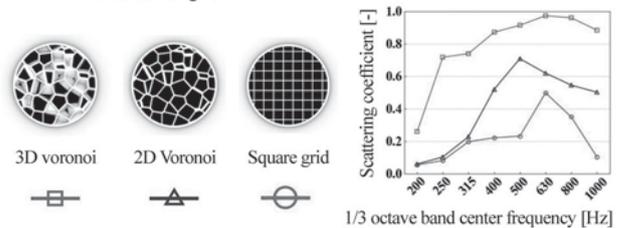


Fig.1: The relationship between the scattering coefficient and the sample shapes. (3D Voronoi, 2D Voronoi, and Square grid)

3-9-10

3-9-10 マイクロスリットを用いた音響メタサーフェスの吸音特性に3Dプリンタの造形精度が与える影響

The effect of 3D printer's manufacturing accuracy on sound absorption of acoustic metasurface using microslits

☆木村祐希, 奥園健(神戸大院・工)

- ◆音響メタサーフェス吸音体 (AMS 吸音体) を 3D プリンタで造形した場合、その造形精度に起因する試験体の寸法誤差が、理論・数値解析値と実測値との乖離の原因となることが指摘されている。
- ◆本研究では、マイクロスリットを用いた AMS 吸音体 (SAMS 吸音体) を対象に、3D プリンタの造形精度に起因する試験体の寸法誤差が理論値と実測値の対応に与える影響を検証した。
- ◆結果、SAMS 吸音体では肉厚、スリット幅に寸法誤差が生じやすいことが明らかとなり、設計寸法の代わりに顕微鏡・ノギスで測定した実寸法を理論解析に用いることで、理論値と実測値の一致が大きく向上することが示された。



Fig.1: A sample of acoustic metasurface using microslits.

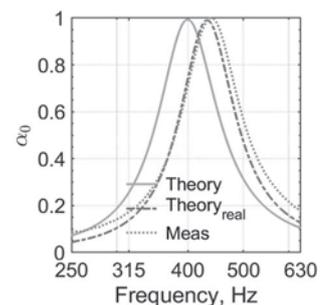


Fig.2: Comparison of theoretical and measured normal incidence sound absorption coefficients.

3-9-11

3-9-11 衝撃力特性の違いによる乾式二重床の床衝撃音レベル低減量の計測と考察

Measurement and consideration of reduction of transmitted impact sound of dry-type floating floors with different impact force characteristics

○富田隆太(日大・理工)

- ◆ Fig. 1, 2 を見ると、タッピングマシン、タイヤ20 cm からの落下も、ゴムボールの床衝撃音レベル低減量とは異なる。ゴムボールは63 Hz 帯域ではバングマシン、タイヤ20 cm 落下とほぼ同様な値を示すが、125~250 Hz 帯域ではタッピングマシンとほぼ同様な値を示すことも重要な視点であると考えられる。
- ◆ 加振周波数特性に相当するコンクリートスラブ素面の床衝撃音レベルで、床衝撃音レベル低減量を補正した結果、タイヤ20 cm 落下の結果に着目すると、63 Hz 帯域から 500 Hz 帯域の範囲で、バングマシンとほぼ同様な周波数特性の傾向を示した。

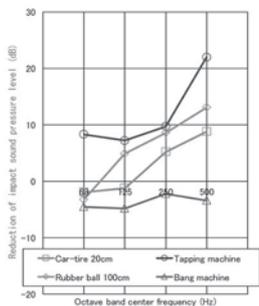


Fig. 1 Reduction of transmitted floor impact sound level for dry-type floating floors B

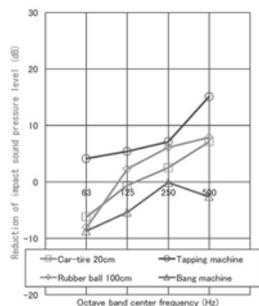


Fig. 2 Reduction of transmitted floor impact sound level for dry-type floating floors C

3-9-13

3-9-13 パルス加振法による防振量の動的ばね定数に関する実験的検討

Experimental study on dynamic spring constant of anti-vibration tatami using pulse excitation method

☆奥山雄太(日大院・理工), 富田隆太, 岡庭拓也(日大・理工)

- ◆ 筆者らは、これまでに防振を考慮した量(以下、防振量)を対象として、床衝撃音と転倒衝突及び歩行時の硬さについて研究を行ってきた [1,2]。
- ◆ 既報^[4]では、歩行感の更なる向上を目指して防振量を試作し、最大衝撃力レベル低減量について検討を行った。
- ◆ 本報では、既報^[4]と同様の防振量を用いて、タイヤ加振時を想定した防振量の動的ばね定数をパルス加振法により計測し、既報^[4]の 63 Hz 帯域における低減量との対応について検討を行った。その結果、タイヤ加振時を想定した防振量の動的ばね定数(振動加速度振幅 15240~24079 cm/s²における動的ばね定数)と 63 Hz 帯域の低減量との対応は、比較的良好な可能性を示した。

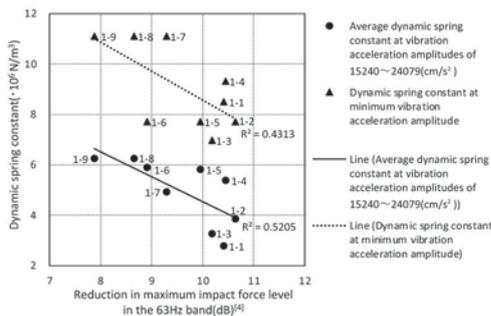


Fig.1: Correspondence between dynamic spring constant and reduction amount in 63 Hz band

3-9-12

3-9-12 防振を考慮したフローリングとカーペットの重量床衝撃音と歩行時の硬さに関する実験的検討

Experimental study on heavy-weight floor impact sound and hardness during walking for flooring and carpet considering vibration isolation

○富田隆太(日大・理工), 奥山雄太(日大院・理工), △阿部今日子(日大・芸術)

- ◆ 筆者らは、既報^[1~3]では、防振を考慮した量を試作し、重量床衝撃音、転倒衝突時の硬さ、歩行時の硬さに関する検討を行った。
- ◆ 本報では、既報^[1~3]を参考に、防振を考慮したフローリング及びカーペットを製作した。その結果、Fig. 1 のケースでは、乾式二重床単体と比較すると、63 Hz 帯域は 5~6 dB の差であった。
- ◆ 歩行感については、F2, C1, C2 は歩きやすいと評価された。

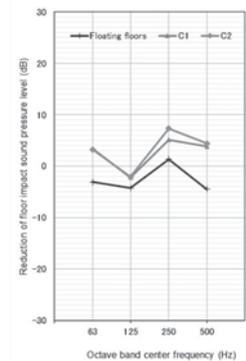
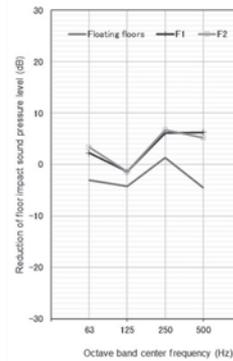


Fig.1: Reduction of transmitted floor impact sound level for floor finishing materials over dry-type floating floors B (Left: floorings, Right: carpets)

3-9-14

3-9-14 床衝撃音を対象とした日用品の落下による衝撃力特性と力積に関する実験的検討

Experimental study on impact force characteristics and impulse caused by falling daily necessities for floor impact sound

○松田貴(防衛基盤整備協会 / 日大・理工), 富田隆太(日大・理工),

△中川英傑(日大院・理工), 岡庭拓也(日大・理工)

- ◆ 床衝撃音遮断性能の評価に用いる標準衝撃源には、タッピングマシン、タイヤ、ゴムボールが規定されている。しかし、これらの標準衝撃源は実衝撃源ではないため、人の生活で発生する衝撃力や周波数特性の全てを網羅していない。床衝撃音遮断性能の評価を標準衝撃源だけでなく、実衝撃源に対応するためには、衝撃力の絶対値や周波数特性を把握することが重要であると考えられる。今後、実衝撃源による床衝撃音遮断性能を示すことが可能になれば、より実生活へ対応した評価が可能になり、音環境の居住性能評価法として非常に有用であると考えられる。
- ◆ 本報では、標準衝撃源 3 種および日用品 24 種の落下による最大衝撃力、衝撃時間、力積等の衝撃力特性を示し、衝撃力暴露レベルおよび周波数特性から標準衝撃源と類似した日用品について検討した。

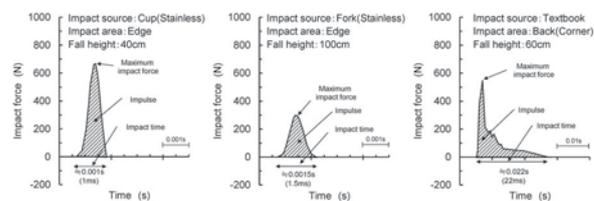


Fig. 1 Example of impact force measurement for daily necessities.

3-9-15

3-9-15 共同住宅において片持ちスラブに配置される居室の床衝撃音遮断性能に関する検討

A study on floor impact sound insulation of rooms arranged on cantilevered slabs in multi-family housings

○黒木 拓, 大脇雅直(熊谷組)

超高層共同住宅では、居室からの眺望の観点から妻住戸の角に柱を配置しない場合がある。このようなプランでは一部が片持ちスラブとなり、床衝撃音遮断性能への影響が懸念される。本報では、片持ちスラブに配置される居室の床衝撃音遮断性能について検討した。

- ◆片持ち梁と通常の大梁ではスラブ端部の拘束力に差はみられない。
- ◆片持ち梁によるスラブの共振による影響は軽微と考えられるが、柱の拘束の影響により、共振の影響が軽微になる可能性が考えられる。
- ◆スラブの共振による影響が軽微な場合には、スラブ素面の予測値と実測値は良く対応しており、片持ちスラブであることの影響はみられなかった。一方、1次固有振動数が63Hz帯域となる場合には柱の影響により共振による影響を受けにくくなることがうかがえる。

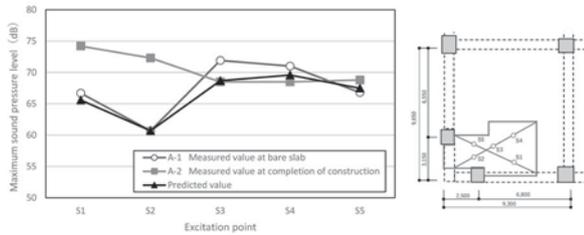


Fig.1 Measured and predicted values of 63 Hz band of heavy weight floor impact sound at each excitation point

3-9-17

3-9-17 重量床衝撃音測定結果に与える騒音計の違いの影響 -バンドパスフィルタの周波数特性および時間応答に着目した検討-

Effects of sound level meter on the measurement of heavy floor impact sound, -Studies focusing on the frequency characteristics and time response of bandpass filters-

○田中ひかり(大成建設 技術センター)

- ◆異なる機種5台の騒音計を用いて、実験室内で重量床衝撃音を同時に測定したところ、測定結果が最大で約4デシベル異なった。
- ◆原因のひとつは騒音計のバンドパスフィルタの周波数特性にあり、通過帯域の外側における周波数特性が影響することがわかった。
- ◆また、低い周波数では騒音計のバンドパスフィルタの時間応答が騒音計によって異なり、その影響も受けることがわかった。
- ◆各騒音計およびそのバンドパスフィルタは規格に則っている。したがって、バンド値を用いた測定を行うのであれば、この程度の差が生じることを認識して測定値を取り扱うべきと考える。

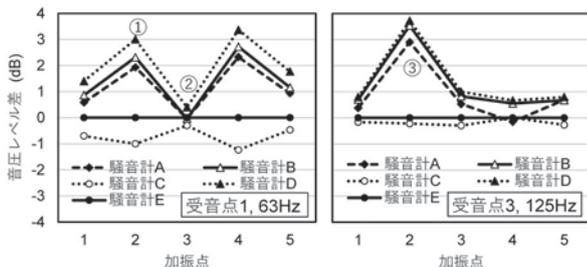


Fig.1: Results of floor impact sound. (Octave band. Difference with result of sound level meter E)

3-9-16

3-9-16 実験室における乾式二重床構造の重量床衝撃音レベルの実測値と解析値の対応

Comparison of measured and numerical results of heavy-weight impact sound pressure level of dry-type double floor system in laboratory

○曹達(東大・工), 會田祐(長谷工技研), 井上尚久(九大・芸工), 佐久間哲哉(東大・工)

- ◆実験室における乾式二重床構造の重量床衝撃音レベル予測手法を構築した。実験室での実測結果と比較し、予測手法の妥当性を確認した。
- ◆スラブ素面: 40 Hz 帯域以下では床スラブ端部の境界条件により解析値は実測値と乖離したが、50 Hz 帯域以上は良く対応した。
- ◆乾式二重床(ボール衝撃): 二重床支持脚の拘束度 R の影響が小さい。接着剤ありの支持脚は線形ばねに近い挙動をした。
- ◆乾式二重床(タイヤ衝撃): 拘束度 R が小さいほど予測値は実測値と対応する。接着剤の効果は小さいと言える。

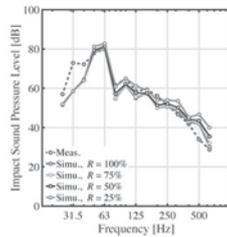


Fig. 1: Measured and simulated impact sound pressure level of double floor system using ball impact.

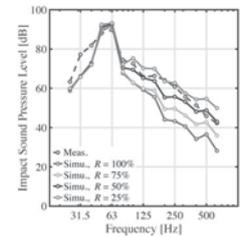


Fig. 2: Measured and simulated impact sound pressure level of double floor system using tyre impact.

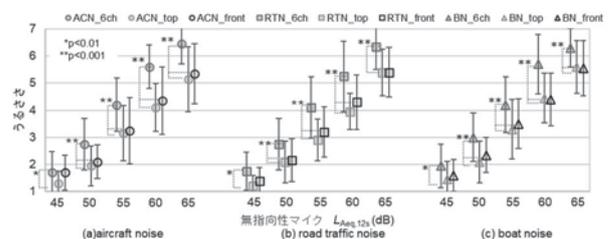
3-10-1

3-10-1 音の到来方向がうるささ評価に及ぼす影響に関する主観評価実験

Subjective experiment on the effect of the direction of noise on noisiness

☆小田切彩夏(東大大学院), 米村美紀(前橋工科大), 森長誠(大同大), 坂本慎一(東大生研)

- ◆2020年より、羽田空港において新飛行経路の運用が開始された。航空機騒音の騒音レベルは幹線道路や街路沿いの住宅街と同様の値となっているにもかかわらず、騒音による苦情が散見されている。
- ◆交通騒音の種類によるうるささの違いには、音の空間性が影響しているのではないかと考え、本研究では、騒音の空間性とうるささ評価の関係に関する被験者実験を行った。
- ◆試験音はピンクノイズの周波数特性を調整して作成し、航空機騒音(ACN)、道路交通騒音(RTN)、船舶騒音(BN)を模した3条件を設定した。提示の方向条件は、6方向条件、上方向条件、正面方向条件の3条件とした。実験参加者には、音のうるささを1~7(数字が大きいほどうるさい)の7段階で評価するよう指示した。
- ◆全方位から騒音にさらされる状況では、ある特定の方向から騒音にさらされる場合よりもうるささ感が増す傾向が示唆された。



3-10-2

3-10-2 自動車交通騒音の主観評価における低周波数成分の影響

Influence of low frequency components on subjective evaluation of road traffic noise

☆藪内 和真(神奈川大院・工学研), 横島 潤紀(神奈川県環境科学センター/神奈川大・工), 森長 誠(大同大・建築), 牧野 康一, 土肥 哲也, 横山 栄, 小林 知尋(小林理研), 山崎 徹(神奈川大・工)

- ◆著者らは、低周波数成分を十分に含む自動車交通騒音を対象とした主観評価実験から得られたデータに構造方程式モデリングを適用し、評価構造モデル (Fig.1) を構築した。得られたモデルから、低周波数成分の有無が不快感に及ぼす影響を考察した。
- ◆モデルより、低周波数成分の有無は不快感に対し、圧迫感・振動感を經由し、間接的に影響を及ぼすことが確かめられた。このことから、一定以上のレベルの低周波数成分の騒音の暴露が圧迫感・振動感を引き起こし、不快感を強めるといった、低周波数成分の不快感に対する間接的な影響が考えられる。したがって、低周波数成分を含む騒音の暴露が与える心理的影響を検討する際、圧迫感・振動感を考慮することが重要であることが示唆される。

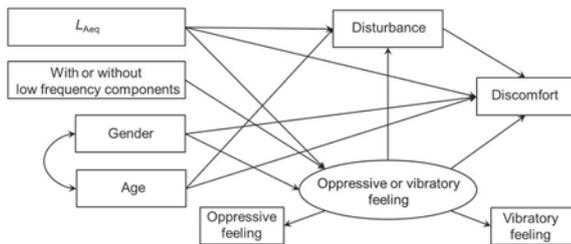


Fig.1 Evaluation structure model for road traffic noise

3-10-4

3-10-4 生活環境中の設備機器騒音の聴感評価実験

Subjective evaluation test of equipment noise in living environments

○米村美紀(前橋工科大), 菅原彬子(近畿大), 坂本慎一(東大生研)

- ◆設備機器騒音には卓越した狭帯域成分が含まれることがあるが、このような騒音の評価に A 特性音圧レベルが適しているかは検討が求められている。一方、近年はより精緻な評価指標であるラウドネスレベルが規格 (ISO532 シリーズ) として整備され、種々の騒音評価への使用が期待されている。
- ◆そこで、騒音評価指標と心理反応の関係を調べることを目的とし、生活環境中で録音した、狭帯域成分を含む騒音を対象に大きさ・うるささの評価実験を行った。
- ◆両指標とも主観評価値とは良い対応を示したが、ラウドネスレベルの方がやや当てはまりが良い結果であった。

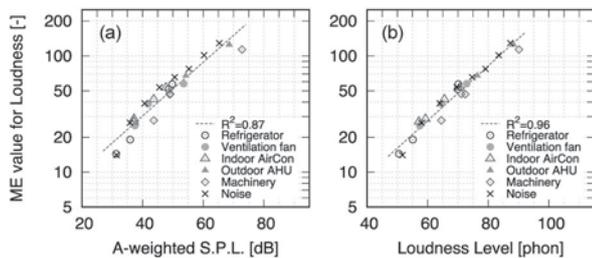


Fig.1: Relationship between Subjective Rating of Loudness and Indicators for Environmental Noise. (a) A-w. S.P.L., (b) Loudness Level.

3-10-3

3-10-3 道路交通騒音の曝露と睡眠妨害反応との関係の経年変化

Temporal change of relationship between the exposure and self-reported sleep disturbance associated with road traffic noise

○横島潤紀(神奈川県環境科学センター/神奈川大学) 森長 誠(大同大学) 山崎 徹(神奈川大学)

筆者らは、国内の幹線道路近傍で実施された社会調査により得られたデータセットを用いて、道路交通騒音の曝露反応関係の二次分析を行っている。本稿では、戸建住宅の居住者を対象に、既報における高度のアノイアンス反応に引き続き、騒音曝露と自己申告に基づく睡眠妨害反応との関係の経年変化を検討した。ロジスティック回帰分析の結果に基づき、夜間等価騒音レベル (NL) と高度の睡眠妨害の反応割合 (%HSD) との関係の経年変化を構築した (下図)。前報の結果に基づき、時間帯補正等価騒音レベルと高度のアノイアンスの反応割合との関係と比較すると、比較的高レベルの騒音に曝露されている場合にのみ、高度の睡眠妨害の割合が低減していることを明らかにした。

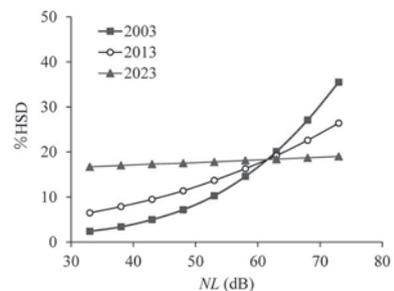


Fig.: Relationship between night-time exposure level and percentage highly sleep disturbed people by survey year

3-10-5

3-10-5 発話レベルを媒介変数とした騒音レベルによる飛沫放出量の予測モデルに関する研究

Study on models for predicting droplet emission due to noise level by using speech level as a mediating variable

☆中山晴登(茨城大院), 石塚崇, △山本ミゲル(清水建設技研), △長谷部弥(清水建設技研), 辻村壮平(茨城大院)

- ◆周囲の音圧レベルから、人の発話行動特性を考慮した飛沫の発生系パラメータの提案に向けて、クリーンルームで発話実験を実施した。
- ◆A 特性音圧レベルを説明変数、発話レベルを目的変数とした非線形回帰分析を行い、周囲の音圧レベルと発話レベルの関係式を得た。
- ◆エアロゾルと飛沫の度数は、Gamma-Poisson 分布に従うことから、一般化線形混合モデルを想定した飛沫放出量の予測手法を検討した。
- ◆発話レベルと飛沫放出量の関係を把握するために、ポアソン回帰分析を実施し、発話レベルとエアロゾル及び飛沫放出量の関係式を得た。
- ◆上記の二つの関係式を用いて、発話レベルを媒介変数とした A 特性音圧レベルによるエアロゾル及び飛沫放出量の予測計算法を導出した (Fig.1)。

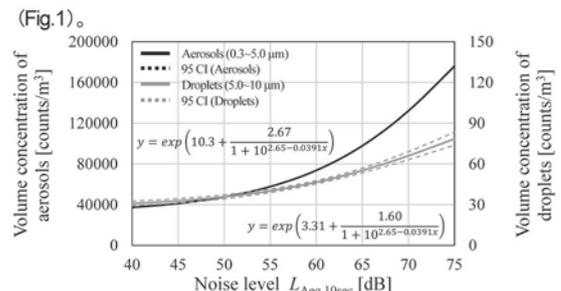


Fig.1: Relationship between A-weighted sound pressure level and volume concentration of aerosols and droplets

3-10-6

3-10-6 博物館における音刺激の抽出と印象評価 —センサリーマップ作成に向けた検討—

Extraction of Sound Stimuli and Impression Evaluations in Museums
—Research towards Creating Sensory-friendly Map—

☆木下晏里沙, 上野佳奈子(明治大),
△小内美空, △高瀬雄士, △吉澤望(東京理科大)

- ◆センサリーマップ作成に関わる検討として、東京国立博物館にて、特徴的な環境要素の抽出及び定型的な印象評価を行った。
- ◆特徴的な環境要素の抽出結果からは、ASD 当事者と一般参加者(学生)の評価で概ね共通の着眼点が見られた。定型的な印象評価の結果においては、感覚特性による評価の差が確認された(Fig.1)。

Table 1: Characteristics of the stimulating sound environment

Element	Characteristics of physical quantity	Characteristics of sound source, environment
Noise	High SPL	• Lots of human traffic, Multiple sources
	Fluctuation of SPL	• Nonstationary Noise ex.) air curtains, announcement
	BGN with characteristic timbres	• Air-conditioning sound is the main sound source, with characteristic timbres
	Other	• Noises from outside of the rooms
Reverberation	Long reverberation time (approx. 2 seconds or more)	• High ceilings, Large space • Reverberate • Mixture of various noise

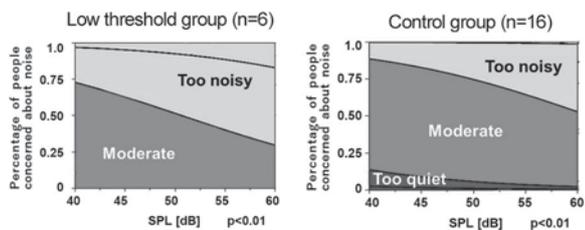


Fig.1: Results of ordinal logistic regression analysis

3-11-1

3-11-1

A Comparative Study of Tikhonov Regularization Parameter Selection Approaches for the Pressure Matching Method

☆周桐(龍谷大院), 安枝和哉(東京医療保健大), 片岡章俊(龍谷大)

- ◆本研究では多点音圧制御法(Pressure Matching)による複数音場再生における適切な Tikhonov 正則化パラメータ β の決定手法を比較する。 β は伝達関数に依存し、合成音場の性能と安定性を左右するため、適切な値を設定する必要がある。
- ◆本研究では、L-curve、一般化交差検証(GCV)、および系統的探索(SS: Systematic search)を従来法とし、探索効率を向上するため二分探索(BS: Binary search)で β を決定する手法を提案した。シミュレーション実験から、提案手法は同等な性能を提供したうえで、適切な β を決定するための試行回数が最大 N から $\log_2 N$ に向上した。

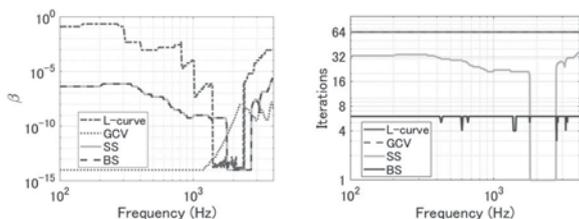


Figure 1. Determination of regularization parameter β and corresponding iteration requirements in each method

3-10-7

3-10-7 構造方程式モデリングを用いた オープンプラン型オフィスの音環境と 知的生産性の関係に関する基礎的検討

A Basic Study on the Relationship between Sound Environment and Intellectual Productivity in Open-Plan Office Using Structural Equation Modeling

☆中橋 樹香(近畿大院), △富樫 建五(大和ハウス工業), 原田 和典(岡山県立大), 菅原 彬子, △長澤 康弘, 平栗 靖浩(近畿大), △岩切 幸伸(ココロ)

- ◆オープンプラン型オフィスでの実地調査により実環境下での音環境と作業者の印象評価を収集し、音環境と印象評価との関係について、構造方程式モデリングを用いた検討を行った。
- ◆個人作業では騒音レベルは集中しやすさに与える影響が最も大きい結果となった。
- ◆複数人作業ではある程度のノイズが作業者の生産性を向上させるといったことが示唆された(Fig.1)。

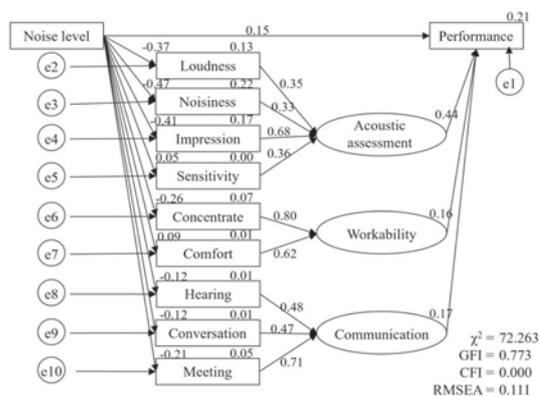


Fig. 1: Structural Equation Modeling (Multi-person work)

3-11-2

3-11-2 ラメ関数による非円形アレイ形状の半径 定式化を用いた音場生成に関する検討 Study on sound field creation using radius formulation of noncircular array with Lamé function

☆坂口 智弘(早大理工), 及川 靖広(早大理工)

- ◆球面調和関数を用いた音場制御の研究では球体アレイや円形アレイが多く用いられている。しかし多くの部屋の形状は矩形であり、空間を有効活用できていない。
- ◆本研究ではラメ関数を用いることで立方体の表面上の点を定式化し、矩形アレイにおける球面調和関数を用いた平面波の生成を行った。
- ◆シミュレーションの結果、矩形アレイにおいても球体アレイと類似する結果が得られ、本検討の妥当性が示された。

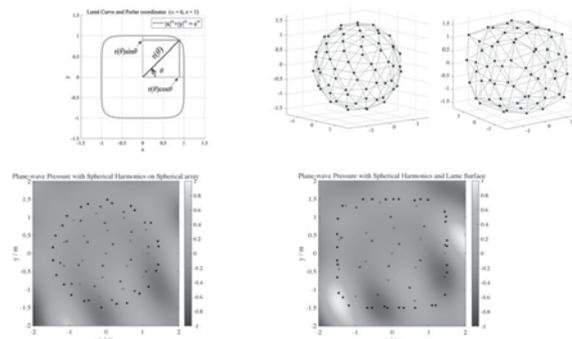


表-1 アレイ内側に対する目的音場からの相対誤差の平均

球体アレイ	ラメ関数アレイ
0.29	0.35

3-11-3

3-11-3 矢状面/横断面配置・左右チャンネル独立型トランスオーラルシステム

Channel independent transaural system using sagittal/transverse-plane-arranged loudspeakers.

☆ 伊藤大智, 鮫島俊哉 (九大・芸工)

- ◆ トランスオーラルシステムは、バイノーラル信号を聴取者の耳元に、スピーカを用いて厳密に再現する立体音響再生手法である。
- ◆ Fig.1 は、提案するシステムのスピーカ配置およびブロック図である。
- ◆ Fig.2・3 は、従来の方式に比べ、提案方式が聴取者の前後左右の移動に対するロバスト性が非常に高い方式であることを示す。

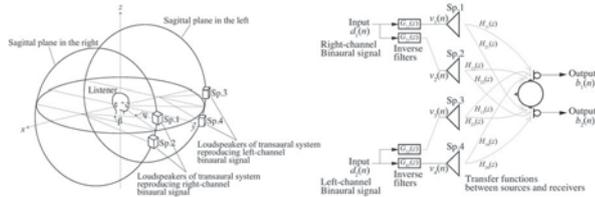


Fig. 1 : Proposed transaural system. Left: Loudspeaker arrangement. Right: Block diagram.

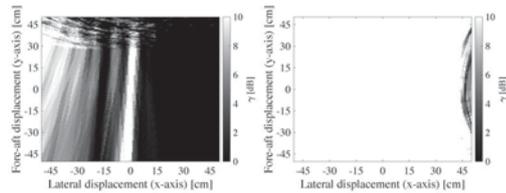


Fig. 2 : Cross-talk cancellation γ in the right ear for listener's lateral and fore-aft movement. Left: Stereo dipole. Right: Proposed system.

3-11-5

3-11-5 Predicting Acoustic Directivity of Home Appliances Using Gaussian Models

Tawhidul Islam Khan (Saga University), ☆ Muhammad Syafril Hidayat (Saga University)

- ◆ In this article, we propose experimental results in an anechoic chamber to analyze the influence of loud noises that are generated from home appliances, as vacuum cleaner and hairdryer.
- ◆ In this experiment, the sound propagation of each home appliance has been observed, followed by combination state to mimic real conditions that possibly appear.
- ◆ Results of the experiments have been compared with the results of calculations for each home appliance individually as well as combinedly to predict the directivity of the generated sound field. The result of the experiment confirmed with higher accuracy on each condition of the combinations.
- ◆ One sample result (Fig.1) of SPL in prediction and measurement for vacuum cleaner and hairdryer is added.

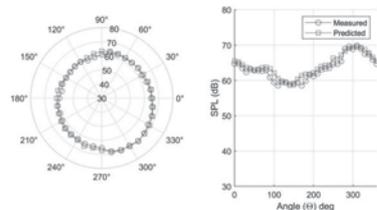


Fig.1: Directivity sound field for combination of home appliances at 1.6 kHz

3-11-4

3-11-4 DNN 音場画像雑音除去における波数スペクトル損失を考慮した高周波音場への拡張

Extension of deep sound-field denoiser to high-frequency sound fields considering wavenumber spectral loss

☆ 岸田一馬(早大理工), 谷川理佐子(NTT/早大理工), 石川憲治(NTT), 及川靖広(早大理工)

◎偏光高速度干渉計を用いた光学的音響計測

- ・偏光高速度干渉計によって2次元音場の定量計測が可能である
- ・音を瞬時に可視化計測し、音場イメージング画像を記録できる
- ・計測された画像には光学的・電気的な雑音が多く混入してしまう

◎波数スペクトル損失を考慮した雑音除去

- ・近年 DNN を用いた音場画像の雑音除去が提案された
- ・高周波音場での高い推論精度を目的として、新たに Fig.1 に示す波数スペクトルを考慮した手法を提案
- ・従来手法よりも高い推論精度を得ることができた

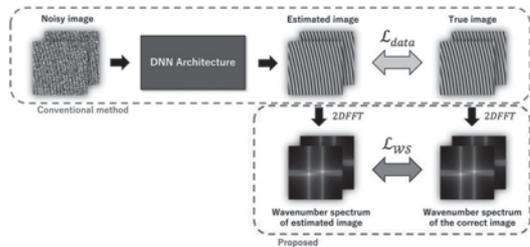


Fig.1 Schematic diagram of the proposed method

3-11-6

3-11-6 音響信号を用いた時系列解析によるモーターの故障予知に向けた基礎検討 -機械学習を用いた小型 DC モーターの過電圧故障予知に向けた一検討-

Basic study for prediction of motor failures by time series analysis using acoustic signals - A study on prediction of overvoltage failures of small DC motors using machine learning -

☆ 池垣 聡司, 坂本真一, △ 平山智士 (滋賀県立大)

- ◆ 音響信号を用いた時系列解析によるモーターの故障予知に向けた検討として、小型 DC モーターの過電圧故障予知を検討した。
- ◆ コンデンサマイクで録音したモーターの駆動音データ 33 分を、学習データとして機械学習を行った。(Fig. 2, Fig. 3)
- ◆ 自己相関係数の特徴量とした結果、10 分前での故障予知を確認した。

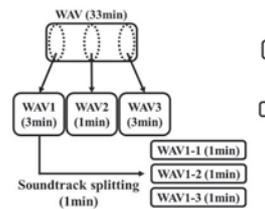


Fig. 2 Recording data analysis procedure (33 min).

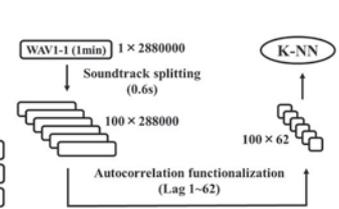


Fig. 3 Recording data analysis procedure (1 min).

3-11-7

3-11-7 グラフ信号処理を用いたマルチチャンネル参照信号予測のチャンネル削減に関する検討

A Study on Channel Reduction of Multi-Channel Reference Signal Prediction Using Graph Signal Processing

© 小塚詩穂里, 伊藤弘章, 鎌土記良 (NTT)

- ◆ 広帯域の騒音を能動的に抑圧する手法として Feedforward 型のアクティブノイズコントロール (ANC) が広く用いられている。
- ◆ 3次元のANCを行うために多数のマイクによる信号予測を行うことは有望だが実応用上少数マイクに絞る必要がある
- ◆ 多数のマイクから一つの信号を作るのは非効率的、多数の信号を作るべき
- ◆ マイクを選択し非選択マイクの信号を同時に補間する手法としてグラフ信号処理の活用が有用である可能性がある

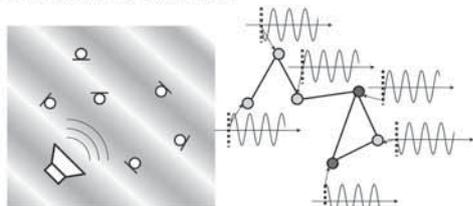


Fig. 1 実環境におけるマイク配置と音響信号におけるマイクのグラフ化

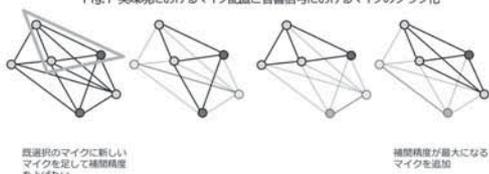


Fig. 2 マイク選択イメージ

3-11-9

3-11-9 2次系の常時推定法に関する予備的検討

A preliminary study on methods of always estimating secondary path

© 藤井健作 (コダウェイ研), 大塚安実治 (関西大), 菅木植史 (千葉工大)

Filtered-x LMS 法を導く評価関数が成立するための条件として、2次系の推定に誤差がある場合に構成される擬似最適解を、誤差がない場合に構成される最適解に近付けることが求められる。これは両解の差を小さくすることが Filtered-x LMS 法を安定に動作させるための条件となることを意味する。更に、2次系フィルタの係数に重畳する推定誤差は確率変数と考える必要がある。このことは、推定誤差の大きさだけでは騒音制御の安定性が保証されないことを意味する。

Filtered-x LMS 法による安定した騒音制御は、2次系を常時推定して2次系フィルタの係数の更新を継続することによる平均化効果によって実現される。但し、2次系推定誤差は発散の確率が1/2未満となる大きさに維持されることが求められる。そのためには、推定誤差が所要の大きさとなることが保証される方法で2次系は推定される必要がある。

実用時には、2次系の推定処理と騒音制御を連続して動作させて場合、システムの起動から騒音低減効果が実感できるまでの待ち時間が苦痛と感じられないように短くすることも重要である。本報告では、2次系推定誤差の大きさを規定する所要値を段階的に小さくする操作を追加することで待ち時間が短縮されること、その推定に際して制御されるステップサイズを2次系インパルス応答の形状に合わせることで待ち時間が更に短縮されること、この両操作を組み合わせることで待ち時間が1分以内に短縮されることを示す。

3-11-8

3-11-8 Shared-error NLMS アルゴリズムによる音響エコー雑音キャンセラの可変ステップサイズ調整法とそのダブルトーク環境下における性能評価

Experimental evaluation under double talk condition of acoustic echo and noise canceller using shared-error NLMS algorithm with variable step size

© 岩居健太, 西浦敬信 (立命館大)

本稿では、可変ステップサイズ Shared-error NLMS (VSS-SENLMS) アルゴリズムによる音響エコーキャンセラ (AENC) (Fig. 1) のダブルトーク環境下における性能評価に関して報告する。本 AENC は音響エコーキャンセラ (AEC) と適応雑音キャンセラ (ANC) を SENLMS アルゴリズムにより同時最適化する。また、可変ステップサイズを導入することで、非正常環境下でも音響エコーおよび背景雑音の抑圧性能を保持できる。本稿では、未検討であるダブルトーク環境下における本システムの性能評価を実施する。計算機実験の結果、ダブルトーク環境下で本システムの性能が大幅に劣化し、その原因が分散の推定方法にあることが明らかとなった。一方、分散を高精度に推定できた場合、ダブルトーク検知器なしで音響エコーと背景雑音の抑圧性能が保持されることが明らかとなった (Fig. 2)。

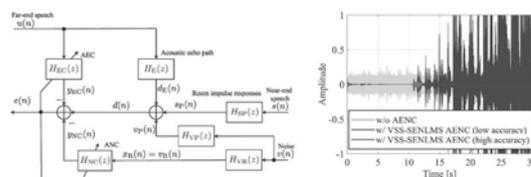


Fig. 1 SENLMSに基づくAENC

Fig. 2 残差信号の時間波形

3-P-1

日本語母語話者と学習者による感情表現の韻律的特徴

—持続時間と F0 曲線に着目して—

Prosodic features of emotional expressions by native Japanese speakers and learners: Focusing on duration and F0 contour

© ☆王可心 (神戸大院), 林良子 (神戸大)

- ◆ 本研究では、日本語母語話者と5か国語を母語とする学習者を対象に、「喜び」「怒り」「悲しみ」「驚き」の4種類の感情で発された短文を分析し、日本語学習者の韻律習得における問題点を示した。
- ◆ 分析の結果、中国語話者とイタリア語話者は比較的に問題が少なかったが、韓国語話者とベトナム語話者はやや問題があり、ロシア語話者は持続時間も F0 も日本語話者と異なる特徴が見られた。
- ◆ Fig. 1 と 2 は「なに」の発話文の持続時間と F0 曲線を示す。

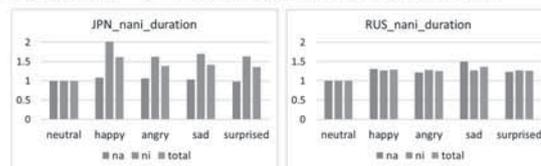


Fig. 1: Variation in the duration of emotional speech "nani" by native Japanese speakers (left) and Russian speakers (right)

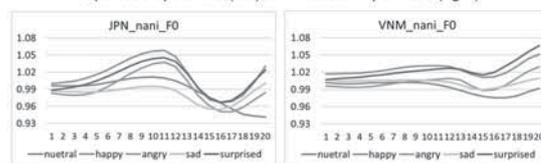


Fig. 2: Variation in the F0 contour of emotional speech "nani" by native Japanese speakers (left) and Vietnamese speakers (right)

3-P-2

3-P-2 中国語を母語とする上級日本語学習者の拗音における /o/ /u/ の生成

The production of /o/ /u/ in palatalized consonants by advanced Japanese learners whose native language is Chinese.

○井下田貴子(早大), 饗庭絵里子(電通大), 荒井隆行(上智大)

- ◆背景: 中国語を母語とする日本語学習者は、/o/ /u/ を後続母音とする拗音の発音に誤りが生じやすく、上級でも習得が困難なケースがある。先行研究では /o/ → /u/ に誤る傾向があると報告されている。
- ◆本研究の目的: 中国語を母語とする上級日本語学習者の発話した2母音 /o/ /u/ の混同の程度について明らかにする。
- ◆調査対象者: 中国語を母語とする上級日本語学習者6名(女性)
- ◆分析: Praatにより拗音の子音に後続する母音のF1, F2を抽出し、判別分析を行った。
- ◆結果: /u/ の判別率は83%, /o/ の判別率は89%で、先行研究と異なる結果となった。

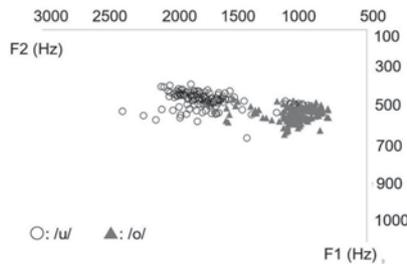


Fig.1: Vowel space of /o/ /u/. (○: /u/, △: /o/)

3-P-4

3-P-4 日本語の調音-音声データベースの拡張

Expansion of the Japanese articulation and acoustics database

☆池田恵里, 若宮幸平, 吉永幸靖(九州大), 桂田浩一(東京理科大)

- ◆日本語の調音-音声データベースの公開を目指し、データベースの拡張を行った。
- ◆拡張したデータは、男性2名と女性8名の計10名分の調音-音声同期データである。
- ◆調音運動の観測には3次元磁気センサシステム(3D-EMA)を使用し、調音観測点は舌上の3点と、両唇、下顎とした。
- ◆発話文はATR503文を使用。

- ◆データベース内容
 - 音声データ
 - 3D-EMAによる座標データ
 - EGGによる声帯振動データ
 - Julius-speech/segmentation-kitによる自動ラベリングデータ



Fig.1: 測定点を●で示している。

3-P-3

3-P-3 日本語促音対立の知覚における先行母音長の役割: 発話速度の影響

Role of preceding vowel duration on perception of geminate consonant contrast in Japanese: Influence of speaking rate

○田嶋圭一(法政大・文), 横谷有紀(法政大院・人文科学研)

- ◆目的: 促音の有無を判断する知覚の手がかりとしての先行母音長の役割の大きさが、キャリア文の発話速度の影響を受けるかを検討した。
- ◆方法: 「この文字は「ハット」/「ハト」と読みます」という音声を提供し、「ハット」と「ハト」のどちらに聞こえたか回答を求めた。キャリア文の発話速度3(速い, 普通, 遅い; 参加者内) × ターゲット語の先行母音長3(0.7倍, 1.0倍, 1.4倍; 参加者内) × ターゲット語の閉鎖区間長6(15ms刻み; 参加者内) 日本語を母語とする大学生14名の回答を分析した。
- ◆結果: 発話速度が速いほど、また先行母音なが長いほど促音であると知覚されやすかった。有意な交互作用はなく、先行母音長の役割の大きさはキャリア文の発話速度の影響を受けなかった。

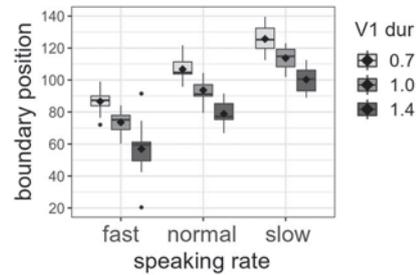


Fig. 1 Category boundary position as a function of speaking rate and preceding vowel (V1) duration. Diamonds indicate means.

3-P-5

3-P-5 痛みを表すオノマトペ表現とその音声特徴 —アンケート調査を基にした認識分析—

Onomatopoeic Expressions of Pain and Their Phonetic Features —A survey-based cognitive analysis—

○横山安紀子(日本大学), 浅野恵子(順天堂大・医)

- ◆医療現場において、痛みを表現する方法には、視覚的アナログ(VAS)、数値的評価(NRS)、表情尺度(FRS)等のスケールがある。
- ◆日本語では痛み表現のオノマトペが多用され、約8割の患者がそれで自分の痛みを理解されたと認識している。
- ◆医療従事者と患者は表現語彙について共通認識があるか不明である。
- ◆調査方法: 痛みの表現の一般的認識について、20代から70代の日本語話者180名にアンケート調査を行った。「痛み」を表すオノマトペ語彙44種について、身体感覚的及び認知的評価を7件法で、また経験する「痛み」について身体部位と回数を8件法により集計した。
- ◆結果とまとめ: 痛みを表す代表的および、形態の異なるオノマトペについて、痛みの度合い、範囲、持続時間、回数(頻度)、深さ、鋭さ、圧迫感の7項目についての回答を分析した。特殊拍をもつ形態のオノマトペはCVCVの反復形(例: ズキズキ)よりも、「鋭さ」を除く6項目において低い評価平均値を示す傾向となった。また、全体的として女性は男性よりも各項目で高い評価値を示す傾向が見られた。



Fig. 1 Onomatopoeis related to Zuki-zuki.



Fig. 2 Onomatopoeis related to Hiri-hiri.

3-P-6

Analysis of paralinguistic characteristics in spontaneous Japanese and Burmese speech: Focused on the Japanese “n” and the Burmese “n”

☆ Lae Lae HTUN, Tetsuya SHIMAMURA (Saitama Univ., Grad. Sch. Sci. Eng.)

Mee SONU (Saitama Univ., Grad. Sch. Hum. Soc. Sci.)

Abstract

This study investigates how “n” sounds convey paralinguistic information (PI) in spontaneous Japanese and Burmese speech. Analysis reveals distinct F0 dynamic patterns between the languages. Japanese “n” exhibits four patterns (flat, fall, gradual fall, and rise), while Burmese has three (flat, fall, and gradual fall). Notably, the gradual rise pattern is specific to Japanese and functions as a form of backchanneling or acknowledgment in response to the topic of dislikes expressed by another person. Additionally, Burmese speech samples have shorter durations and a higher pitch in general compared to Japanese in all patterns.

Keywords: Paralinguistic Information (PI), Burmese, Japanese, spontaneous conversations, F0 characteristics

3-P-8

3-P-8 混合音声から抽出した発話の聴取に関する聴者・難聴者の比較 Comparison of Normal Hearing and Deaf/Hard-of-Hearing Subjects on Evaluation of Personal Speech Enhancement

◎藤江 匠汰 安 啓一(筑波技大) 小林 彰夫(大和大)

- ろう・難聴者の聴取における課題
 - ろう・難聴者は、複数話者の混合音から目的音の聴取が困難
 - 聴力損失によって目的音の手がかりが得にくい(選択的聴取)
- 音源分離による抽出音声の主観評価
 - 音源分離手法を用いてろう・難聴者の抱える課題を解決
 - 目的音のSNRによる、ろう・難聴者の聞き取りやすさの変化
 - 聴者による評価との比較
- 実験
 - 音源分離: SpeakerBeam (JNAS 64.8時間で学習)
 - 評価データ: 話者2名によるATRバランス文の読み上げ
千葉大学3人会話コーパスをSNR=0, 10, 20 dBとして混合
 - 主観評価: ろう・難聴/聴者それぞれ11名によるlistening effort評価
- 実験結果
 - ろう・難聴者、聴者ともに音源分離の前後で有意差が認められた
 - SNRが0 dB, 10 dBのときに有意差が認められた
 - 音源分離手法は、SNRが低い場合にろう・難聴者の聞き取りやすさの改善に役立つと考えられる

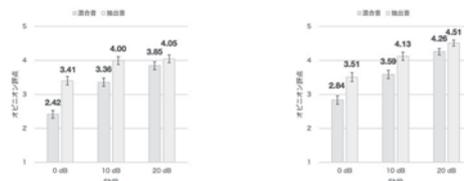


図1: オピニオン評点(ろう・難聴者) 図2: オピニオン評点(聴者)

3-P-10

3-P-10 臨床データを用いた難聴者の語音聴取 Identification of monosyllables by hearing impaired listeners

○杉山由希子(慶應大), △鈴木大介, △新田清一(済生会宇都宮病院)

- ◆ 済生会宇都宮病院耳鼻咽喉科補聴器外来を受診した難聴者(37名)の語音明瞭度検査(67-S語表)の結果を報告する。
- ◆ 今回得られた予備調査の結果を参考にして、将来的には日本語に特化した補聴器調整法の開発を目指す。

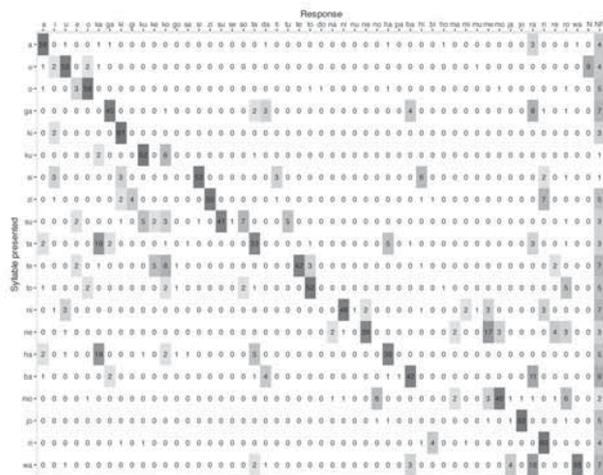


Fig.1: Confusion matrix

The y-axis indicates the syllables presented to the listeners and the x-axis the listeners' responses.

3-P-11

3-P-11 認知症高齢者を対象とする被験者実験のための感情制御音声刺激に関する検討

A study on emotion-controlled speech stimuli for subject experiments with dementia elderlies

◎中西亮太(名古屋市立大院), 齊藤千晶(認知症介護研), 中村篤(名古屋市立大院)

認知症者は、疾患の早期から感情内容の理解に困難をともなう、との報告があるが、感情の強さの違いによる影響は検討されておらず、知見は限定的である。そこで、本研究では、感情表現に関わる韻律的特徴を調整することで、精緻に感情の強さが制御された音声刺激を作成し、認知症高齢者の感情認知能力を推し量るための音声聴取被験者実験への適用を目指す。音声をもつ韻律パラメータを2通りの方法で変化させて、高精度な感情変換音声刺激を作成し、健常な学生を対象とした予備実験で妥当性を検証した。その結果、一方向変換より双方向変換の方が正答率は高くなり、変換時に発生するスペクトルのひずみ等の影響が軽減されることが示された。ただし、正答率があまり高くならなかった感情対もあり、高精度な音声刺激の作成方法にはさらなる検討の余地がある。

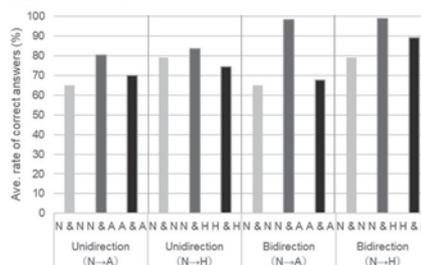


Fig.1: Average rate of correct answers by conversion method and emotion pair.

3-P-12

3-P-12 認識結果の信頼度推定手法を利用した高齢者の加齢による衰え推定の可能性

Estimation of Age-Dependent Degradation Using Estimation reliability of speech recognition results

☆飯田篤央, 脇田由実(大阪工大院), △松田千登勢(摂南大)

- ◆従来、高齢者の日常会話音声の基本周波数の時間変化特徴を用いて、加齢に伴う高齢者の理解力の衰えを推定する手法を研究しているが、高齢者の会話音声認識率が施設利用者と自宅生活者とは大きく差があることから、音声認識結果の特徴から高齢者の衰え推定を行うことを試みた。
- ◆入力波形に僅かな加工(発声始端を微量に変更・雑音を微量に重量)を施し、加工前後の音声認識結果の変化率から音声認識結果の信頼度を推定する手法を提案している(Fig.1)。この手法を用いて、施設利用高齢者と自宅生活高齢者との認識性能の差が推定できれば、そこから衰え度を推定できる可能性がある。
- ◆認識実験を行った結果、施設利用高齢者の認識結果変化率の方が高い傾向にあり(Fig.2)、本信頼度推定手法を高齢者の衰え度合の推定に利用できる可能性が見受けられた。

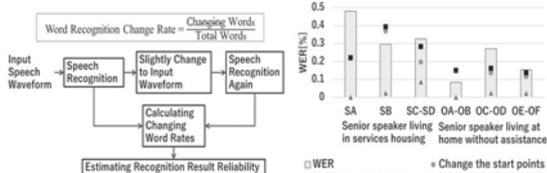


Fig.1 Process for estimating the reliability of speech recognition results

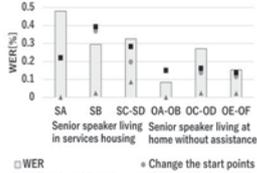


Fig.2 Relationship between the misrecognition rate and the rate of change in elderly conversations.

3-P-14

3-P-14 Convergent 型声帯物理モデルを用いた非定常流下の声門内圧測定実験

Experimental measurement of intra-glottal pressure under unsteady flow using a convergent-type vocal fold physical model

☆井上晴貴, △久米田員男, 船戸優希, 徳田功(立命館大)

- ◆アクリル製 M5 声帯物理モデル (Fig.1) を作成し、収束角を 10°、声門幅を 0.24mm に固定した。
- ◆定常流下および流量を正弦波状に変動させた非定常流下で、声門下圧と声門内圧を測定した。
- ◆非定常流下実験における声門下圧の時間変化と、声門内圧差(定常流下の声門内圧から非定常流下の声門内圧を引いたもの)の関係を調べた。
- ◆上流 (x=50.8mm) と下流 (x=52.83mm) で、定常流下の声門内圧と非定常流下の声門内圧の大小関係が逆転することが分かった (Fig. 2(a), (b))

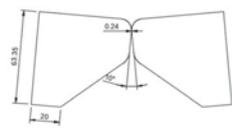


Fig.1 M5 model

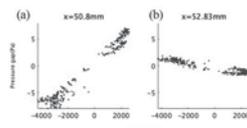


Fig.2(a) Pressure gap at x=50.80mm, (b) Pressure gap at x=52.83mm

3-P-13

3-P-13 アバター支援言語学習システムを用いた質問回答訓練とリピーティング訓練による英語表現定着度の比較評価

Retention effects of form-focused English speaking training based on question-answering and repeating tasks using avatar-assisted language learning system

☆野村采加(同志社大院・理工学研究科), 田中龍弥(同志社大院・理工学研究科), 山本誠一(同志社大・研究開発推進機構), 加藤恒夫(同志社大・理工)

- ◆当研究室では、2体のロボットの英会話に学習者1名が参加して英会話の対面訓練を行うロボット支援言語学習システムを提案してきた。
- ◆遠隔訓練を可能にするため、コンピュータに表示されるアバターと対話を行う遠隔学習形式のアバター支援言語学習システムを開発した。
- ◆学習者を質問回答グループとリピーティンググループの2つに分け、学習定着度の比較を行った。さらに、従来使用してきた対面形式のロボット支援言語学習システムと遠隔形式のアバター支援言語学習システムで学習定着度の比較を行った。
- ◆質問回答グループとリピーティンググループの比較では、質問回答グループの方が学習定着度は高く、ロボット支援言語学習システムとアバター支援言語学習システムの比較では学習定着度の差はなかった。

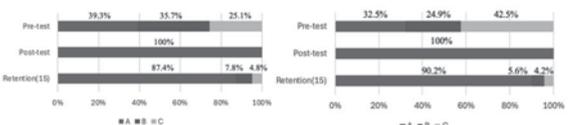


Fig.1: Retention test results for response tasks that participants in QA Group achieved A scores in the post test (avatar on the left, robot on the right)

3-P-15

3-P-15 声帯膜下部の溝および発声条件が声帯振動に与える影響

Influence of sulcus of the vocal membrane on the vocal fold vibration

☆廣野凌, 小畑大樹, △上松日, △松原大輔(立命館大), △西村剛(大阪大), 徳田功(立命館大)

- ◆声帯膜下に溝が存在する物理モデルを用いて吹鳴実験を行った。溝の曲率半径が3種類のモデルおよび、内転角度を3通り設定し、声帯膜のみの振動から声帯膜-声帯振動に遷移する条件を分析した。
- ◆吹鳴実験を通じて音声および高速度動画を計測した。
- ◆声帯膜の独立性の高さが声帯膜振動の発生に寄与していることが確認できた。また、独立性の低いモデルでも内転を加えることで声帯膜のみの振動が発生することが分かった。

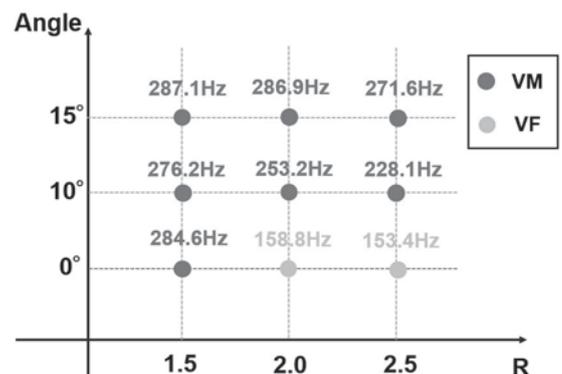


Fig.1 Fundamental frequency and vibration type of each model and condition.

3-P-16

3-P-16 アカゲザルの声帯物理モデル実験: 声帯膜の有無が声帯振動に与える影響について

Experimental study on physical model of the vocal folds for rhesus macaques: Effect of vocal fold membranes on the vocal fold vibration

○山本康旭, 藤江真己人, 杉江昂輝(立命館大), △西村剛(大阪大), 徳田功(立命館大)

- ◆アカゲザルの喉頭をマクロCT撮像し、生理データに基づく物理モデルを用いて、吹鳴実験を行った。
- ◆声帯膜がある物理モデルと声帯膜を切除した後の物理モデルに対して、声門下圧、音声、高速動画を計測し、比較を行った。
- ◆摘出喉頭においても同様の実験を行い、比較を行った。
- ◆声帯膜が存在しないことにより、オンセット圧は上昇し、発声効率は下降した。
- ◆摘出喉頭でも同様の傾向がみられ、声帯膜は大きな声を出すのに使用されているものと示唆された。

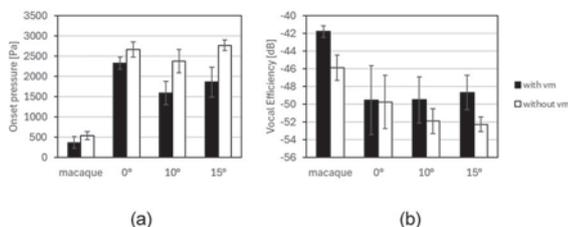


Fig.1: Result of Onset pressure(a) and vocal efficiency(b) under each condition

3-P-18

3-P-18 チューブ発声前後の声道形状の変化が声帯物理モデルに与える影響

Effect of tube-induced change in tract shape on physical model of the vocal folds

☆足立恭平(立命館大), △水本直宏(立命館大), 座喜味絃武(立命館大), 徳田功(立命館大)

- ◆チューブ発声による声道形状の変化が発声に及ぼす影響を調べた。
- ◆MRI型声帯物理モデルを用いて、発声閾値声門下圧および声門上圧、発声効率、高速動画を計測した
- ◆チューブ発声前および発声中と比べ、チューブ発声後に、発声効率が上昇することが分かった。

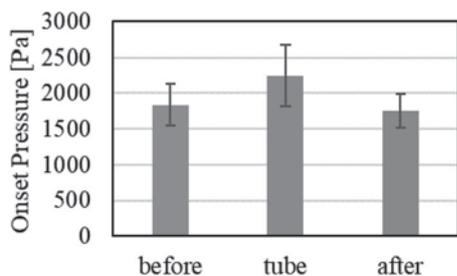


Fig.1: Comparison of vocal efficiency between vocal tract models before, during (tube), and after tube phonation.

3-P-17

3-P-17 声帯物理モデルを用いた母音生成実験

Generation of vowel sounds using a physical model of the vocal folds

☆小畑大樹, 廣野凌, △上松日, △松原大輔(立命館大), △西村剛(大阪大), 徳田功(立命館大)

- ◆声帯物理モデルと声道物理モデルを組み合わせて、母音生成実験を行った。
- ◆声帯物理モデルには、二種類(声帯膜を持たないヒト型と、声帯膜を持つ非ヒト型)を用いた。
- ◆声道が存在する場合、非ヒト型声帯モデルの振動が安定することが確認された。
- ◆非ヒト型声帯モデルに対して流量を動的にし変化させたところ、モデルによっては、短時間で100Hz程度のピッチ変動が確認された。

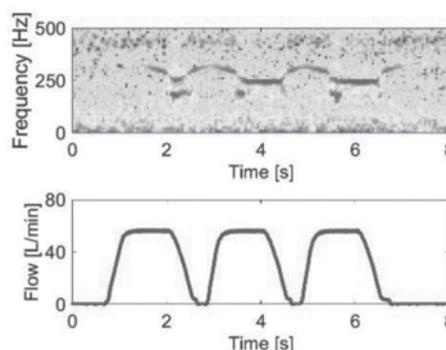


Fig. 1 Spectrogram and flow rate of r2.0 model

3-P-19

3-P-19 ゲームアプリを用いたチューブ発声訓練支援システムの評価

Evaluation of Tube phonation Support System using a Game Application.

☆村井武人(甲南大院), 北村達也(甲南大), 川村直子(姫獨大)

研究の背景と目的

- ◆本研究は、チューブ発声を用いた音声治療の研究である。
- ◆チューブ発声には少なくとも1500回の反復が必要であるため、患者が安定して訓練できる方策が求められている。
- ◆そこで、シリアスゲームを導入した発声訓練支援システムを開発し、高齢者を対象にチューブ発声訓練に対する有効性を調査した。

実験結果

- ◆実験ではアンケートと発声時間を集計した。Fig.1は参加者ごとの発声時間を1日の平均でまとめたものであり、実線がシステムを利用した実験群、点線がシステムを利用せずに訓練した統制群である。
- ◆アンケートの結果では、システムは楽しい訓練の実現や、訓練に対するモチベーションの維持に対して有効であることが確認できた。
- ◆Fig.1より両群の発声時間に大きな差があり、システムを用いることで、指示した発声時間(5秒間)で安定して訓練できると考えられる。

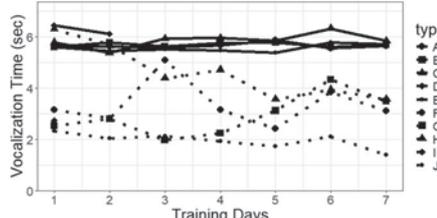


Fig.1: Averaged duration of tube phonation

3-P-20

3-P-20 音声トレーニングに向けた発話運動 MRI を利用したスマホアプリの開発

Development of a smartphone application for voice training using speech motor MRI

○高野佐代子・土田義郎(金沢工大)

Sayoko TAKANO and Yoshio TSUCHIDA (Kanazawa Institute of Technology)

- ◆音声トレーニングにおいて、発話時の舌運動に興味を持ってもらうことは重要である。
- ◆その一環として、スマートフォン（以下スマホ）で収録した指定のセリフの音声をもとに、簡磁気共鳴画像 MRI (Magnetic Resonance Imaging) の発話運動の動画をスマホ上で確認できるシステム「Voice Tuning Helper」を試作したので報告する。



Figure 1. The interface of the application for a smartphone

謝辞) 本研究のシステムは丸山翼氏(株式会社プロノハーツ)が製作したものである。国立国語研究所が開発した「リアルタイム MRI 調音運動データベースを利用している。心より御礼申し上げます。

3-P-22

3-P-22 glimpse model を用いた雑音中のポップアウトボイスの検出

Application on the glimpse model for pop-out voice detection in noise

○河原 英紀(和歌山大), 天野 成昭(愛知淑徳大), 北原 真冬(上智大) 牧 勝弘(愛知淑徳大), 坂野 秀樹(名城大)

- ◆の雑音の中でも目立って聴こえる声(ポップアウトボイス)を、聴覚フィルタバンクを用いて分析し、Glimpse model を参考にして性質を調べた。
- ◆Glimpse で形成されるマスクの占める領域の割合を独立変数とするよりも、領域での突出量を2帯域(高/低)で集計した量を用いる方が高い相関を示した。

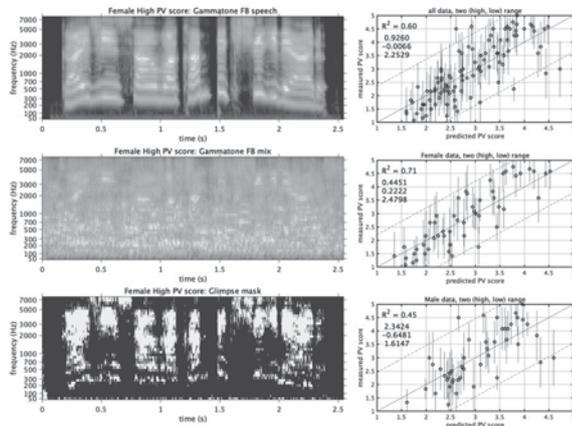


Fig.1: Left: The auditory spectrogram of speech and mixed-signal and the Glimpse mask (a female with a high-Pop-out Voice score). Right: Scatter plots of predicted and measured PV scores (all talkers, female and male).

3-P-21

3-P-21 非対面での会話難易度と対人評価にリアルタイム声質変換が与える影響: 協力ゲームを用いた調査

The effect of real-time voice conversion on conversion difficulty and interpersonal evaluation at non-face-to-face: Investigation using of cooperative game

☆菅野聖真, 松井淑恵(豊橋技科大院)

- ◆リアルタイム声質変換 (VC) を介した際の簡単なコミュニケーションに対する評価実験を実施した。
- ◆自身の音声がVCへ入力されることを事前に伝えた協力者と、実験設定を何も伝えない参加者(男性10人、女性10人)間で簡単なオンラインゲームを実施させ、参加者が協力者に対して感じたコミュニケーションの難易度と対人魅力を質問紙で評価させた。
- ◆VCを介した条件とVCを介さない条件間で、コミュニケーション難易度の評価に有意差は見られなかった。対人魅力のうち、好意や関心の対象としての評価である「情緒的魅力尺度」にも有意差は見られなかったが、課題の遂行手段としての評価である「道具的魅力尺度」に有意差が見られた (Fig. 1)。

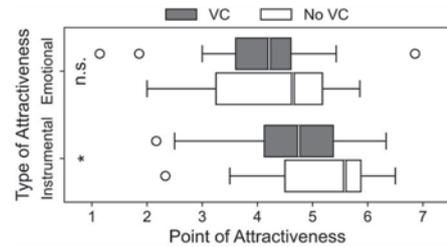


Fig.1: The evaluation result of interpersonal attractiveness. (white vertical line: median, open circle: outlier, asterisk: p < 0.05)

3-P-23

3-P-23 謝罪の音声に対する誠意の評価: 人間話者と TTS 合成音声の比較

Evaluation of sincerity for speech of apology: a comparison of human speakers and TTS synthesized speech

☆澤田拓郎, 松井淑恵(豊橋技科大院)

- ◆TTS 合成音声人間話者の謝罪音声と同等に、相手に謝罪の誠意を伝えられるかどうかを検討した。また、両者の謝罪音声の韻律を操作し、謝罪から受ける誠意に影響を及ぼすかどうかを調査した。
- ◆人間話者4名(男性2名、女性2名)、TTS 合成音声4種(男性2種、女性2種)の謝罪音声「申し訳ございません」を原音声として使用した。原音声の抑揚と発話速度を操作し、音声刺激を作成した。
- ◆参加者17名(男性14名、女性3名)は各謝罪音声に対して許容できるかどうかを7段階で評価した。
- ◆話者条件(人間話者・TTS 合成音声)で謝罪によって伝わる誠意に差があること、話者条件によって、音声の抑揚と発話速度が謝罪音声の誠意評価に与える影響が異なることが明らかになった。

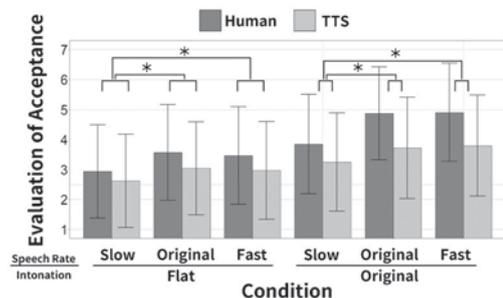


Fig.1: Mean and SD of acceptance.

3-P-24

聴覚的不快感を決定づける音響統計量

Sound statistics for auditory discomfort

☆丸山玄徳(静岡社会健康医学大学院大, 日本学術振興会), △矢部伶治, △糸井知美, △本吉勇(東大)

わたしたちは、黒板を爪で引っ掻く音などに対して非常に強い不快感をもつ。もし任意の自然音の聴覚的な質感を全般的に説明できる特徴空間において不快音と関連する/しない情報を判別することができれば、それは聴覚的不快感を決定づける音響特徴に関する普遍的な知見となるだろう。そのような特徴空間の候補に、近年提案されている二段階スペクトル空間がある。これは、音波の時間微細構造周波数に関する一次スペクトルと、時間変調とスペクトル変調からなる二次スペクトルで構成され、広範囲の自然音の質感知覚を高精度に記述することができる。本研究では、明らかに不快な音を含む広範囲の自然音に対する不快度の評価データを実験から集め、不快度評価と高く相関する二段階スペクトル空間における帯域を同定することを試みた。分析により、非不快音のスペクトルはより低い変調周波数にパワーが集中しており、不快音とは異なるスペクトルの傾向があることが示された。この結果は、一次のみならず二次まで含めたスペクトルが、不快感の要因として重要である可能性を示唆している。この特定された「不快領域」が不快感をもたらすのは、「不快領域」が自然界の多くの音に共通するスペクトル特性から逸脱している点にあると考えられる。

3-P-25

メロディとコード進行の自動生成を用いた聴覚マスキングに基づく歯科治療音の快音化

Comfortable sound design based on auditory masking with automatic generations of melody and chord progression

☆林拓哉(阪産大院), 高橋徹(阪産大), 西浦敬信(立命館大), 中山雅人(阪産大)

- ◆従来の歯科治療音の快音化手法では、コード進行による制御音はあまり快適な音にならないという問題があった。これは、コード進行だけでは楽曲の要素が足りないためだと考えた。そこで、本研究ではメロディとコード進行の自動生成を用いた聴覚マスキングに基づく歯科治療音の快音化手法を提案する。
- ◆Fig. 1に提案手法の概要図を示す。提案手法では、歯科治療音のピーク周波数からコード進行とメロディを生成するためのキーを決定し、決定したキーとコード進行のルールよりコード進行音とメロディ音を生成し、制御音として歯科治療音と同時に受聴することで快音化を行う。

Fig. 2に快音度の主観評価実験の結果を示す。3名による評価の結果より提案手法 (PM: Proposed Method) では、従来手法 (CM: Conventional Method) よりも1.0ポイント快音度が向上していることが確認できた。

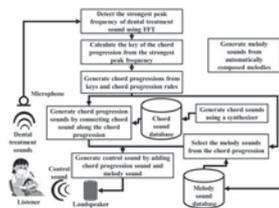


Fig.1 Overview of the proposed method

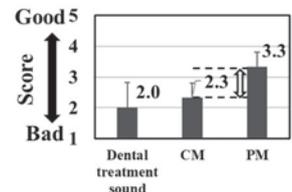


Fig.2 Mean opinion score with five grades for comfortable

3-P-26

3-P-26 事象関連電位によるEV走行音の注意機能評価

-高周波帯域での影響調査-

Evaluating Attention to EV Driving Sounds Using Event-Related Potentials -Impact of High-Frequency Band-

☆村上寛名, 亀山勇希, 石光俊介(広島市大),

- ◆本研究では順応及び選択的注意に着目し、脳波の事象関連電位から走行騒音に対する注意機能を定量することを目的とし、前報から1000 Hzの音源を新たに加えた4音源での実験を行った。
- ◆P3の成分得点では音圧レベル増加時、変動する周波数帯域が高域であるほど有意に大きくなる傾向。
- ◆P3潜時では音圧レベル減少時、増加時ともに変動する周波数帯域が高域であるほど有意に短縮する傾向。
- ◆今後の検討として、刺激音の時定数を考慮することで実際の走行音に近づけた実験を行う。

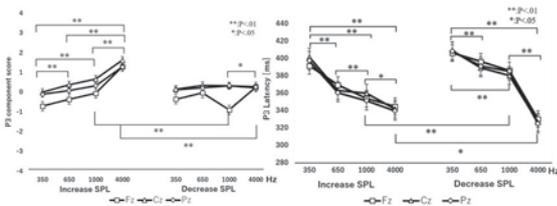


Fig.1:P3 component score and latency of SPL increased and decreased in the frequency bands with center frequencies of 350, 650,1000 and 4000

3-P-27

3-P-27 左右耳への外因的注意が耳音響放射に与える影響

Effect of exogenous attention to left and right ear on otoacoustic emissions

☆松毛円佳, 石坂勇毅(千葉大), 山岸慎平(NTT CS 研), 藤平晴奈(九州大), 大塚翔(千葉大), 古川茂人(静岡社会健康医学大学院大/静岡県立総合病院/NTT CS 研), 中川誠司(千葉大)

- ◆片方の耳に対して内因的に注意を向けると、聴覚末梢の応答である耳音響放射 (Otoacoustic emissions: OAEs) が増強することが報告されている。これは、左右耳の感度がトップダウンに調整され得ることを示唆している。
- ◆一方、呈示された刺激の方向に強制的に注意が誘導される場合もある。これは外因的注意と呼ばれるが、その効果は、時間経過とともに弱まり、対側耳へ注意が向けられるようになる (Inhibition of return: IOR)。
- ◆本研究では、外因的注意、及び、IOR が OAE に与える影響を調べることで、ボトムアップに左右耳の感度が調整されている可能性を検証した。
- ◆手がかり刺激 (Cue) の提示直後は、外因的注意によって Cue と同側耳の標的的刺激 (Target) に対する反応時間が減少し、時間が経過すると、IOR によって対側耳での反応時間が減少することが確認された。(Fig. 1A)
- ◆一方で、Target によって誘発された OAE は、統計的な有意性は確認されなかったものの、Cue と Target の時間間隔が短い条件では、大半の実験参加者において Cue と同側耳の OAE が増強された。(Fig. 1B) 外因的注意の影響は受けるが IOR の影響は受けにくい可能性がある。今後はこの仮説を精査したい。

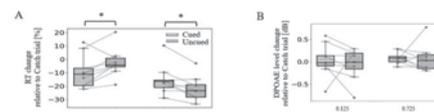


Fig. 1 Changes of reaction time (A) and DPOAE level (B) relative to Catch trials. Individual data are plotted as gray lines. *p<0.05 (simple main effect).

3-P-28

3-P-28 距離方向に対する聴覚的注意の方位依存性の検討

Directional selectivity of auditory selective attention in depth direction.

☆田中優希(熊本大), 寺岡諒(室蘭工大), △寺本渉(熊本大)

- ◆本研究では、注意を向ける方向が距離に対する聴覚的注意効果に及ぼす影響について検討した。
- ◆特定の距離に注意を向けさせない条件(注意なし条件)と、標的音の呈示確率を調整して、32 cm (near), 160 cm (far) のいずれかに注意がひきつけられるよう仕向けた条件(注意あり条件)から構成され、それぞれ前方、側方、後方の3方向で実施された。
- ◆全ての方向条件において、注意を向けた位置で最も標的音に対する感度が高く、標的音と競合音の弁別が他の距離に比べてよくできていることを示した。
- ◆方向条件の効果はみられず、距離に対する注意効果に方位依存性が無い結果となった。しかし、標的音を競合音と誤認した確率である誤警報率の概形が方向により異なることは興味深い。

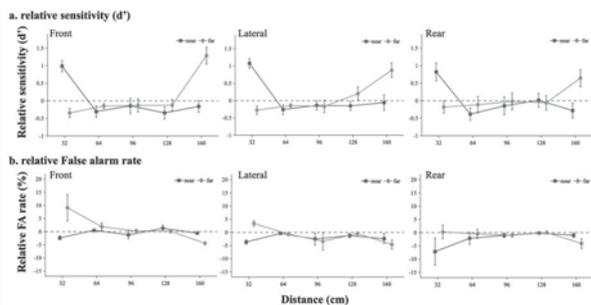


Fig.1: Relative results (each attention condition - control condition) as function of distance of the target sound sources. a. relative sensitivity for the target sound, b. relative false alarm rate for the target sound. Error bars denote the standard error of the mean.

3-P-30

3-P-30 感情評定の異方性に対する視覚情報の影響

Visual information and negative emotional bias for sounds coming from behind

○川島 尊之(帝京平成大)

- ◆音が喚起する快・不快、覚醒感のような感情は、音が前方から提示されるときと、後方から提示されるときで異なるという報告がある。例えば後方ではより不快になり、より覚醒の度合いが高くなる。
- ◆これは聴覚が後方を含めた全周囲への警戒を担っていることと関連していると考えられている。
- ◆こうした前後の方位についての異方性が、聴覚以外の感覚モダリティとどのように関係するかについての研究報告を著者は見つけていない。本稿では視覚情報と感情評定の異方性の関係を研究した。
- ◆実験ではスピーカを含めた後方を視認可能なよう前方に鏡を配置した。実験ブースの照明の有無により視覚情報を利用可能な条件と、利用が難しい条件を設け、その間で感情評定を比較した。
- ◆実験データの探索的な分析からは、異方性への視覚情報の影響の有無について判断することが難しいこと、視覚情報自体は快・不快の感情評定に影響することがわかった(暗いほうが不快に感じる)。
- ◆今後異方性に関するより詳細な分析を続けたいうえで、異方性と視覚情報の関係について研究する。

3-P-29

3-P-29 指さし動作が視聴覚空間注意に与える影響

The Effect of Pointing Movements on Visual and Auditory Spatial Attention

☆李戦(東北大通研/院情科研), 坂本修一(東北大通研/院情科研), 塩入諭(東北大/総合知センター)

- ◆本研究では、指さし動作が視聴覚空間注意に与える影響を調査した。
- ◆指さし運動中に文字識別課題、または音圧変化検知課題を実施した。各課題は指さし方向と課題の方向の一致不一致(Same-Opposite)、指さしの有無(Move-Static)の条件を組み合わせた4条件にて行った。
- ◆実験の結果、聴覚課題ではいずれの組の正答率にも有意差は見られなかった一方、視覚課題では、MovS条件の正答率がStatS条件より有意に高かった。
- ◆このことから、聴覚注意は指さしには影響を受けないのに対し、視覚注意は指さし方向に向けられると考えられる。

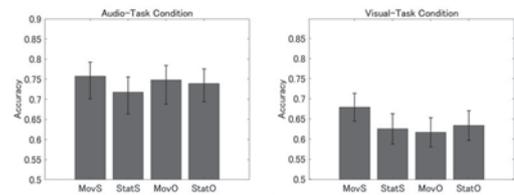


Fig.1: Accuracy of Audio-Task for conditions. (MovS: Move-Same, StatS: Static-Same, MovO: Move-Opposite, StatO: Static-Opposite)

3-P-31

3-P-31 一側性難聴者のための水平面音源定位トレーニング

- 受聴音圧レベルの音源方向依存性の影響 -

Sound source localization training in the horizontal plane for people with unilateral hearing loss - Effects of azimuth angle dependence of listening sound pressure level -

☆稲葉春香(千葉工大・院), 飯田一博(千葉工大・先進工)

- ◆本人のHRTFとヘッドホンを用いて、振幅スペクトルと受聴音圧レベルの音源方向依存性を方向判断の手がかりとして学習できるトレーニングを実施し、音源定位テストでその効果を検証した結果、以下のことが示された。
- ◆擬似一側性難聴者の平均側方角誤差と左右誤判定率は改善し、13日後も維持された。前後誤判定率は改善しなかった(Fig.1)。
- ◆一側性難聴者の平均側方角誤差は改善したが、13日後は維持されなかった。左右誤判定率と前後誤判定率は改善しなかった。
- ◆トレーニングで用いていない男声アナウンスでの平均側方角誤差と左右誤判定率は白色雑音と同程度だった。前後誤判定率はやや悪化した。
- ◆従来のトレーニングと本研究のトレーニングで定位精度の改善率を比較すると、ほとんど差は見られなかった。

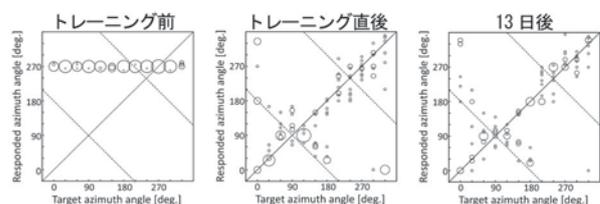


Fig. 1 Responded azimuth angle in the horizontal plane for a normal-hearing subject under simulated unilateral deafness.

3-P-32

3-P-32 耳介形状による上半球正中面の頭部伝達関数のノッチピーク周波数とレベルの推定

Estimation of frequency and level of spectral notches and peaks of the head-related transfer functions in the upper median plane based on pinna shape

☆設楽祐月(千葉工大・院), 飯田一博(千葉工大・先進工)

- ◆上半球正中面7方向(0-180°, 30° 間隔)におけるN/P周波数とレベルを耳介形状により推定した。
- ◆重回帰分析の結果, N/P周波数の自由度調整済み重相関係数はN1で0.60-0.81, N2で0.60-0.73, P1で0.52-0.60, P2で0.66-0.74であった(Fig. 1). N/Pレベルの自由度調整済み重相関係数はN1で0.23-0.72, N2で0.26-0.52, P1で0.43-0.74, P2で0.44-0.70であった(Fig. 2).
- ◆N/P周波数の平均絶対残差はN1で0.05-0.1oct, N2で0.06-0.09oct, P1で0.02-0.03oct, P2で0.09-0.13octであった(Table 1). N/Pレベルの平均絶対残差はN1で3.5-7.2dB, N2で5.8-9.7dB, P1で1.0-1.6dB, P2で2.5-5.0dBであった(Table 2).

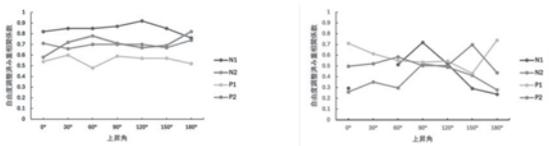


Fig. 1 Adjusted multiple correlation coefficient for N/P frequency

Table 1 Mean absolute residual for N/P frequency estimation [oct]

	N1周波数	N2周波数	P1周波数	P2周波数
0°	0.10	0.09	0.02	0.09
30°	0.08	0.08	0.02	0.13
60°	0.07	0.07	0.02	0.10
90°	0.05	0.06	0.02	0.10
120°	0.06	0.06	0.02	0.10
150°	0.07	0.07	0.02	0.10
180°	0.10	0.07	0.03	0.09

Fig. 2 Adjusted multiple correlation coefficient for N/P level

Table 2 Mean absolute residual for N/P level estimation [dB]

	N1レベル	N2レベル	P1レベル	P2レベル
0°	6.5	6.3	1.6	4.3
30°	5.6	6.0	1.4	3.7
60°	5.6	5.8	1.3	2.9
90°	3.5	5.3	1.0	2.6
120°	4.0	5.3	1.2	2.5
150°	4.8	6.1	1.5	2.5
180°	7.5	9.7	1.3	5.0

3-P-34

3-P-34 インパルス応答の直接波成分を除く後部残響の振幅重み付けが距離感に及ぼす影響調査

Investigation into the effect of amplitude weighting of late reverberation of impulse response on distance perception

☆森垣詠太, 中山雅人, 高橋徹(阪大院・工)

背景 再生音の距離制御は波面合成など物理的再現が注目されてきた
→再現に多くのスピーカーが必要
DRR (直間比) 制御によるシステムの小型化[江川+, 2021]
→DRR 制御の粒度を粗くし処理の軽量化を目指す

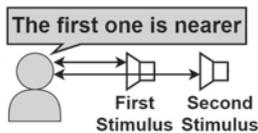


Fig. 1: Experimental condition.

目的 人間の遠近感覚の距離分解能力の確認
DRR 制御のみでどの程度距離感を再現できるのか調査
手段 0.25 (m) 間隔の実 IR (実測のインパルス応答) の音源での遠近判断課題 (Fig. 2)
IR の直接波成分を除く後部残響の振幅を各実 IR の DRR になるよう重み付けした音源での遠近判断課題 (Fig. 3) (Fig. 2, 3: 遠近判断正答率, 横軸第一音, 縦軸第二音)
結果 人間は 0.75 ~ 1.0 (m) 差があれば違いを感じられる DRR だけ再現できれば IR 波形はそれほど重要ではない

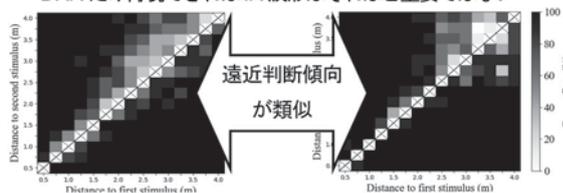


Fig. 2: Perspective perception for IR.

Fig. 3: Perspective perception for controlled IR.

3-P-33

3-P-33 長時間超音波プリバーチャルリード信号による音像定位強調

Augmented Sound Perception Using Long Ultrasonic Pre-virtual-leading Signals

☆今中峻太(立命館大院), 耿毓庭(立命館大), 中山雅人(阪産大), 西浦敬信(立命館大)

- ◆近年, 受聴者に音を高い臨場感で提示可能な高臨場音場再生技術が目ざされており, 心理音響モデルに基づく方式は, 2 ch. のステレオ技術の拡張に基づいている。従来の手法では受聴点における両耳間レベル差や両耳間時間差を再現することで音像提示を行う。
- ◆提案手法では, 先行音効果に着目する。本信号が放射されている間とその前後にかけて先行音効果が生じるような信号を付加することで, 先行音効果を利用した音像定位強調を実現する(Fig. 1)。この際, 付加する信号に可聴域帯の信号を用いると本信号の内容聴取に影響を及ぼすため, 提案手法では, 超音波帯域の信号を使用し, さらに低周波の包絡を与えることで受聴者に知覚されやすくなるよう設計した。
- ◆音像強調方向についての主観評価実験を実施し, 正答率を算出した。結果から, 提案手法の有効性を確認した(Fig. 2)。

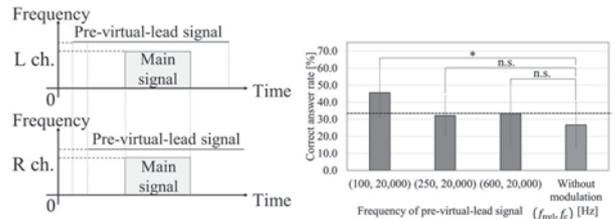


Fig. 1: Concept of the proposed method in time-frequency domain

Fig. 2: Experimental results for direction of sound localization.

3-P-34

3-P-35

3-P-35 骨導ヘッドホンで聴取される音源知覚方向のズレ特性の計測 -広帯域での計測-

Measurement of discrepancy of sound source direction perceived through bone-conduction headphones - pure tones in a whole audible range -

○岩城護(新潟大・工), △杉原晴也(新潟大・工)

- ◆【背景】骨導ヘッドホンでは音源方向は前方(中心)寄りにずれて知覚されることが, 電話帯域(500 Hz から 4000 Hz) の純音に対して報告されている。この特性は骨導ヘッドホンの利点を損なっている。
- ◆【目的】電話帯域よりも広帯域(125 Hz から 16000 Hz) の純音に対して, 骨導ヘッドホンでの音源知覚方向のズレを調査する。
- ◆【手法】両耳間レベル差と両耳間時間差で方向感をつけた純音を気導ヘッドホンと骨導ヘッドホンから提示し, 基準音(骨導音)に対する検査音(気導音)の主観的等価点を計測した。健聴な 20 歳代の 4 人に対してインフォームドコンセントに基づいて実施した。
- ◆【結果】音声帯域外でも, 骨導ヘッドホンで知覚される音源方向は気導ヘッドホンの場合よりも前方(中央)寄りに知覚された。周波数依存性がみられた。

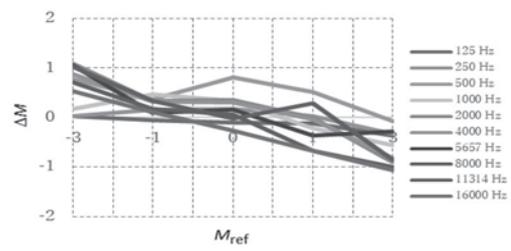


Fig. 1: Discrepancy of perceived direction, ΔM, versus direction of reference sound, Mref.

3-P-36

3-P-36 耳栓着用時の骨伝導音知覚特性に及ぼす 暴露騒音特性の影響： 外耳道内音圧計測による検討

Effects of noise exposure on bone-conducted sound perception during earplugging: Examination by ear canal sound pressure

☆内田知佐, 大塚 翔, 中川誠司(千葉大)

- ◆骨伝導音は耳栓を装着しても容易に知覚できるため、強大騒音下での音声コミュニケーション・デバイスに活用されている。しかしながら、耳栓装着時・強大騒音下の骨伝導音知覚は耳栓の遮蔽特性や騒音特性などの影響を受けて複雑に変化し、メカニズムには不明点が多い
- ◆前報(内田ら, 音認識論(春), 2024)での各暴露騒音条件に対する日本語単語の明瞭度は、無音 > マルチトーカーノイズ > ピンクノイズ > ホワイトノイズとなり、先行研究(無音 > ピンクノイズ > ホワイトノイズ > マルチトーカーノイズ(Osato-Yeboah et al., 2009))とは異なっていた
- ◆本研究では、前報の結果の検証を行うため、骨伝導音声のみを呈示した場合、および強大騒音暴露下で骨伝導音声のみを呈示した場合の外耳道内音圧 (ECSP) を計測した
- ◆得られた ECSP データのスペクトログラムを Fig.1 に示す。マルチトーカーノイズ下では音声情報が多く残存しているのに対し、ホワイトノイズやピンクノイズ下では多くの音声情報が失われていることがわかる。
- ◆ホワイトノイズとピンクノイズは 1000 Hz 以上にパワーが広がるのに対し、マルチトーカーノイズは主に 1000 Hz 以下に多くのパワーをもち、1000 Hz 以上のパワーは低下する。そのため、子音部分の聴取を妨害せず、明瞭度の低下に影響を及ぼさなかったと考えられる。

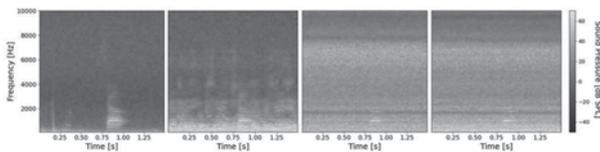


Fig.1 Spectrograms of recorded ECSP of female voice "sya" on 4 conditions (Far left: Female voice only, Left of the middle: Female voice under multi-talker noise, Right of the middle: Female voice under white noise, Far right: Female voice under pink noise).

3-P-38

3-P-38 貼付型振動子を用いた 遠位呈示骨導超音波の検出閾推定

Estimation of hearing threshold of adhesive bone-conducted ultrasound presented to distant parts of the body

☆高橋尚也, 大塚 翔, 中川誠司(千葉大)

- ◆骨伝導で呈示された 20 kHz 以上の高周波音 (骨導超音波) は、頭部から離れた位置に呈示された場合 (遠位呈示) も明瞭に知覚可能である。そのため、体幹や上肢などから振動子に触れた対象者のみに音情報伝達する新型音響デバイスへの応用が図られている。
- ◆一般に骨伝導には振動子装着時に痛みが生じる、見た目が悪いといった装置上の問題があるが、遠位呈示はこれらの問題を解決できる可能性がある。しかしながら、先行研究では伸縮性バンドで体感や上肢に固定しており、実用性の高い振動子固定方法が検討された例はない。
- ◆本研究では、両面テープによって振動子を固定した場合 (粘着方式) の骨導超音波の検出閾を推定し、従来のバンドで押し付けて固定する方式 (従来方式) との比較、および振動子パラメータ (振動子筐体から振動面の突出長、接着面積) と検出閾の関係を評価した。
- ◆振動子を貼付固定できた全ての部位で骨導超音波を知覚可能であり、振動子の突出長や接着面積を増加させることで検出閾を低下させることができた。これらの結果は、貼付型振動子が新型音響デバイスの刺激呈示方式として有用であることを示すものである。



図 1 (a) 振動面がわずかに突出するよう作成された振動子ケース、(b) 振動面、または (c) ケースに両面テープを貼り付けて呈示する

3-P-37

3-P-37 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

Suppression effects of bone-conducted tones by air-conducted pure tones with open ears

☆幸田怜奈(山梨大), 鶴木祐史(JAIST), 小澤賢司, 鳥谷輝樹(山梨大)

- ◆背景: 骨導提示音を用いることで、閉塞耳にイヤホン提示された気導音が約 5 dB 抑圧される (井上ら, 2023).
→ 耳を塞がずに気導音を抑圧できるかは不明。
- ◆目的: 開放耳で気導提示純音を骨導提示純音で抑圧可能か検討。
- ◆方法: 250, 500, 1000, 2000, 4000, 8000 Hz の純音について、気導音を抑圧する最適な骨導音の振幅・位相条件を調整法で決定。得られた条件下で気導提示音の抑圧量を調整法により評価。
- ◆結果: 個人間で相違が見られたが、平均 5 dB の抑圧量が得られた。抑圧効果は周波数に依存 (Fig. 1)。
- ◆考察: 1000 Hz と 4000 Hz においては参加者間で概ね一貫した高い抑圧量が得られた。これは主に伝達される周波数帯域が互いに異なる複数の経路の影響が考えられる。
- ◆結論: 開放耳においても閉塞耳を用いた先行研究と同等の抑圧量となることが明らかとなった。

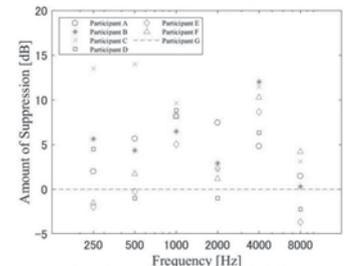


Figure 1: Amount of suppression for AC-tone with BC stimulation ($-20\log_{10}(A_{amp}/A_{50})$).

3-P-39

3-P-39 自然および人為攪乱による環境変化が コウモリの採餌行動におよぼす影響

Changes in bat foraging behavior due to natural and artificial disturbance of the foraging environment - Bat feeding strategy based on sound analysis -

☆平尾碧大, △大西優衣(同志社大), 藤岡慧明(同志社大),

△福井大(東京大), 飛龍志津子(同志社大)

- ◆動物にとって採餌環境の変化に応じた餌や採餌場所の適切な選択は重要な行動である。しかし、野生動物の採餌行動に関しては計測することが難しく、採餌に関する適応行動は不明な点が多い。
- ◆本研究では、エコーロケーション音声より採餌行動を定量的に把握できるモモジロコウモリ (*Myotis macrodactylus*) に着目し、人為攪乱 (池底土砂の排出工事) および自然攪乱 (台風による出水) による餌場環境の変化に対して、モモジロコウモリがどのように採餌行動を変化させているのかを、音響計測を基に調査、分析した。
- ◆その結果、工事や台風による餌場環境の変化は、餌である水生昆虫の数を一時的に減少させ、音響による活動量の計測からは、個々の餌場の環境変化に対応するためコウモリが林内にある複数の餌場を有機的に利用していることがわかった (Fig.1)。

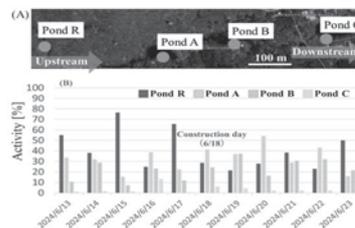


Fig. 1: (A) Multiple ponds used by bats as feeding areas. (B) Proportion of activity in each pond.

3-P-40

3-P-40 飼育環境下と野生におけるつがい期の雄ニホンライチョウの鳴き声の比較

Comparison of the calls of male Lagopus muta japonica during the mating period in captivity and the wild

○石井 要次, 吉田 拓真(積水化学), 土肥 哲也(小林理研), 森川 大輔(富山県立大), △秋葉 由紀(富山市ファミリーパーク)

- ◆絶滅危惧 I B 類に指定されているニホンライチョウの保護活動の一環として、鳴き声を用いた生態系把握に関する研究を進めている。
- ◆前報では複数のニホンライチョウの飼育環境下で鳴き声を長期収録し、鳴き声の種類について考察をした。しかし、野生復帰等の活動を視野に入れたとき、鳴き声の違いが生存率に係わる恐れもある。
- ◆本報告ではつがい期の野生の雄ニホンライチョウの鳴き声の現地収録を行い、飼育環境下の鳴き声との比較を行った。その結果、野生と飼育環境下の鳴き声のバリエーション、周波数や時間的な特徴に差がみられた。



Fig.1: Measurement situation at Mt. Tateyama

3-Q-1

3-Q-1 多重音響イベント検出におけるデータ拡張手法の性能効率性評価と計算負荷の分析

Performance efficiency evaluation and computational load analysis of data extension methods for multiple Sound Event Detection

☆高橋歩夢, 立蔵洋介(静岡大)

- 多重音響イベント検出**は、音声信号のみから同時に発生し得る複数の特定のイベントの発生を検出する技術である。この技術は同時発生する音の分離や検知精度に課題を抱えている。深層学習によって**高精度化が進んだが**、従来手法と比較して**計算リソースの必要量が增大した**。リソースの効率化は計算コストの削減だけでなく、環境に対する負荷の軽減にもつながるため重要であるが、それに反して**リソース効率化に取り組む研究は少ない**。
- 深層学習用データセットは分量に対する収録の負担が大きいため**慢性的に不足**している。従来の精度向上手法ではデータの分量を増やすため、教師データの説明要件の緩和や、信号処理を用いたデータセットの**増強が行われていた**。これらの手法は手間の削減に成功しているが、精度の低下を物量で補っている形であり、**効率的な学習であるとは言えない**。
- 本稿ではデータ拡張手法による**性能向上に対する必要計算コストの変化量の測定**について提案を行う。従来の計算負荷測定指標に加えて、演算中の CPU、メモリ、GPU の各**ハードウェアの使用量の測定**を行う。獲得スコアに対するリソース消費量の割合を計算し、**性能効率性を導出**することで、**効率的にリソースが使用可能な手法の特定が可能**であると考えられる。

3-P-41

3-P-41 深層学習を用いた鳥類の鳴き声の短時間スペクトルの構造の分類

Short-time spectrum classification of bird song using deep learning

☆中谷優太, 山田宏樹, 矢田部浩平(農工大)

- 🐦 鳥類の発声器官である鳴管には、音源 (labia) が 2 つ存在する
- 🐦 キンカチョウなどの一部の鳥類は左右の labia を独立に制御し、異なる音を同時に鳴らす
- 🐦 そのため、鳴き声のスペクトルには1つの調波で表せる構造 (one voice) と表せない構造 (two voice) が含まれる
- 🐦 スペクトルの構造を one voice, two voice, no voice (無音区間) の3つのラベルに分類した

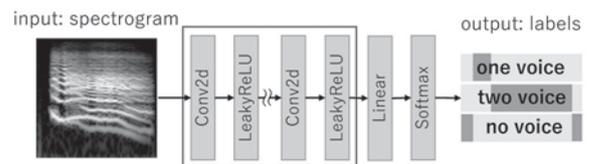


Fig. 1 Network architecture for spectrum classification

3-Q-2

3-Q-2 深層距離学習を用いた話者識別における学習データに対する識別精度の調査

Investigation of recognition accuracy for training data in speaker identification using deep metric learning

☆大島風雅, 中山仁史(広島市大)

- ◆近年、話者認識タスクの解決に深層距離学習で学習した埋め込み空間上で識別性能の高い手法が検討されている
- ◆先行研究では埋め込み空間上での分類を k 近傍法で行っているが、埋め込み空間と同時に学習した深層学習モデルによる分類も可能
- ◆実験から埋め込み空間の学習及び分類モデルの学習に最適な損失の重み係数を調査
- ◆加えて、本手法で高い認識性能を実現するために必要な学習データ量も同時に調査

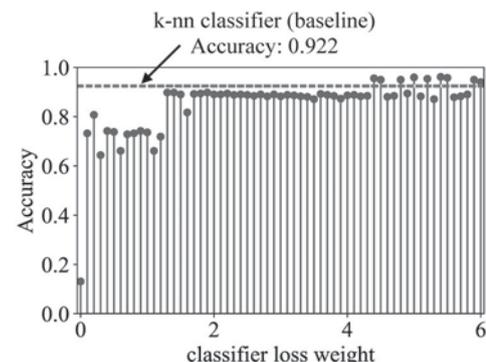


Fig. 1 Recognition accuracy for classifier loss weight, the dash line denotes the accuracy when k-nearest neighbor is used.

3-Q-3

3-Q-3 音源距離推定に音響パワーレベルの 事前知識を活用した音響イベント定位

Sound Source Distance Estimation Utilizing Prior Knowledge of Sound Power Level for Sound Event Localization and Detection

◎佐藤菜緒, 安田昌弘, 齊藤翔一郎, 中山彰 (NTT)

【問題設定】

- ◆ 音響信号を入力として、そこに含まれる音響イベントの種類、開始・終了時刻、及び音源の位置を推定する音響イベント定位タスクにおける距離推定手法を検討。

【提案手法】

- ◆ 室内の音の減衰モデルに基づき音響信号の伝搬を物理モデル化し、本モデルに含まれる一部のパラメータをデータ駆動で推定するハイブリッド型の距離推定手法を提案。

【数値実験と結果】

- ◆ インパルス応答を用いたシミュレーションにより作成した、音響イベント定位データセットによる数値実験を実施。
- ◆ 提案手法により距離推定についての評価指標が19%改善。

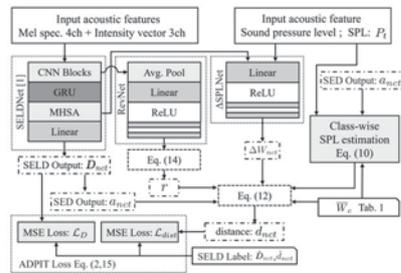


Fig.1: Network architecture of the proposed method.

3-Q-5

3-Q-5 ビジュアルマイクロホンにおける 合焦領域の選別による音質改善

Improvement of Sound Quality in Visual Microphone by Selecting of Focal Area

☆中野隼汰(立命館大院), 歌毓庭, 岩居健太, 西浦敬信(立命館大)

音波により振動する物体をローリングシャッターカメラで撮影し、その撮影動画から音波を抽出する手法(ビジュアルマイクロホン)が提案されている。ここで、撮影動画中に、ぼけた領域(非合焦領域)が存在する場合があります。非合焦領域では、振動物体の変位を高精度に算出することが困難であるため、抽出音波における音質劣化の原因となることが考えられる。しかし、従来のビジュアルマイクロホンによる音波抽出手法では、非合焦領域を考慮せずに音波を抽出する。

本稿では、合焦の度合いを表す「合焦度」に基づいて、合焦領域を選別し、抽出音波の音質改善を試みる(Fig.1)。評価実験では、正弦波を音源として使用した。その結果、提案手法において、一部の周波数に対して抽出音波の音質向上を達成した(Fig.2)。

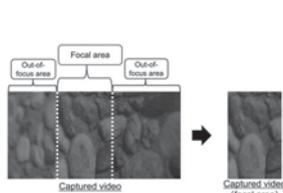


Fig. 1: Selection of focal area.

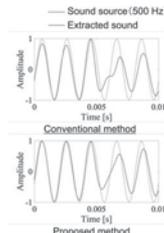


Fig. 2: Waveforms of extracted sound by each method.

3-Q-4

3-Q-4 歌声を含む楽曲における歌声の音高推定 および楽器推定に関する検討

A study on pitch estimation of singing voice and instrument identification for music including singing voice

☆小倉真了, 川凌司, 坂野秀樹(名城大)

- ◆ **自動採譜**: 楽曲から自動的に楽譜を作成すること
- ◆ **音高推定**: 音の高さ(≒基本周波数)を推定すること
 - 自動採譜において最も重要な要素技術
 - 多楽器楽曲における多重音高推定は高難度
- ◆ **楽器推定**: 鳴っている楽器の種類を推定すること
- ◆ 実用的な自動採譜においては、パート譜(楽器ごとに分離した楽譜)の生成が望まれる。
- ◆ 多楽器楽曲におけるパート譜の推定に向けて、深層学習を用いた2つの実験を行う。
- ◆ **実験1**: 歌声を含む楽曲において、**歌声の音高のみ**を推定
 - 同様に各楽器の音高を推定するモデルを作成すればパート譜を推定可能
 - 入力をリアスペクトログラムにした場合とメルスペクトログラムにした場合を比較(データセット: MIR-1K)
 - ◇ **リニア: 70%, メル: 73%** (F1スコア)
 - メルスペクトログラムのほうがより適切と思われる。
- ◆ **実験2**: **多重音高推定と同時に楽器推定**を実行
 - **音高情報と音色情報を分離**することで、多楽器楽曲における多重音高推定の精度を向上させるモデルが提案されている。
 - 音色情報を利用して楽器推定を試みる。

3-Q-6

3-Q-6 拡散モデルを用いた長期 Audio Inpainting の ビジュアルマイクロホンへの適用

Application of Long-Term Audio Inpainting to Visual Microphone with Diffusion Model.

☆大野圭哉, 立蔵洋介(静岡大院・総合科学技術研)

- ◆ **背景**: カメラから音を抽出するビジュアルマイクロホンは抽出原理上周期的な音波の欠損が生じ、復元音の音質を低下させる。
- ◆ **目的**: 長期的なオーディオの欠損信号の推定を行う Audio Inpainting を用いて、ビジュアルマイクロホンの抽出音の欠損区間の補間を実施し、音質が改善されるか調査する。
- ◆ **方法**: カメラからの抽出信号の欠損区間に対し、Audio Inpainting を用いて補間信号を生成し、重みづけ加算を行って接続する。
- ◆ **結果**: カメラからの抽出信号に対してコンテキストと同様の波形の補間を確認し、ノイズが多く重畳する抽出信号の補間が可能であることを確認した。

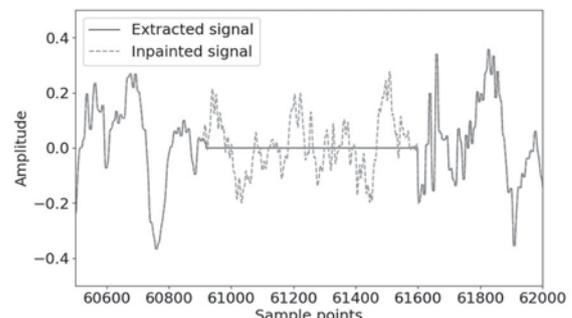


Fig.1: Comparison of extracted signal and inpainted signal waveforms

3-Q-7

3-Q-7 自動車暖機音特性とアノイアンスの関係 第2報 PA値と不協和度からの解析

Relationship between automotive warm-up sound characteristic and annoyance

Second report: Analysis based on PA values and degree of dissonance

☆神原美夏, 石光俊介(広島市大院・情).

△岸川直樹, 藤川智士, 岩田陽明, 松本真典, 菊池正和(マツダ)

- ◆自動車のAWS(早期暖機システム)の暖機音の不快感が問題となっている。第1報ではトーンリティの大きさがアノイアンスに影響することがわかった。そこで本研究では、PA(Psychoacoustic Annoyance)値と不協和度の観点からアノイアンスへの影響を調査した。
- ◆聴感実験では2つのAWS音源とそれらを1音上げた/下げた音源の計6種類を、不協和度の解析ではAWS-1にそれを1,2音上げたものを付加した音源を使用した。
- ◆聴感実験より、Fig.1の順で不快であることがわかる。しかし、PA値(Fig.2)と感覚尺度値や心理音響指標に関連性は見られなかった。
- ◆不協和度については、2音上げたもよりも1音上げたものを付加した音源の方が、不協和度が高かった。これは元音源と1音上げたものの中心周波数が近接していたためと考えられる。

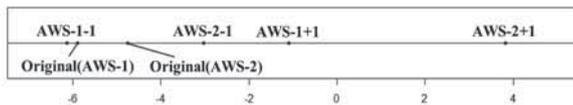


Fig.1: sensory scale value

	Original(AWS-1)	AWS-1+1	AWS-1-1	Original(AWS-2)	AWS-2+1	AWS-2-1
PA値	34.39	34.34	34.48	33.74	35.31	35.00

Fig.2: PA value result

3-Q-9

3-Q-9 骨伝導ヘッドホン使用時の 頭部内振動場と外耳道内部音場の連成数値解析

Vibro-acoustic analysis of human heads and ear canals for bone-conduction headphones

☆尾関航一, 鮫島俊哉(九大・芸工)

- ◆骨伝導ヘッドホン使用時の頭部内振動場について、頭部内振動場と、外部放射音場の連成解析を行った。
- ◆人体頭部を Fig.1 の形状で物理モデル化し、頭部内振動場には FEM を、周囲音場には BEM を適用し、それらの連成数値解析を行う。
- ◆トランスオーラルシステムを人体頭部内振動場に適用し、バイノーラル再生において、骨伝導ヘッドホンを使用した場合でも立体音響再生を可能とすることを目的とする。
- ◆解析によって得られた外耳道内のインパルス応答を再生系伝達関数として、通常の逆フィルタと Acoustic Contrast Filter を用いた制御器を設計した。

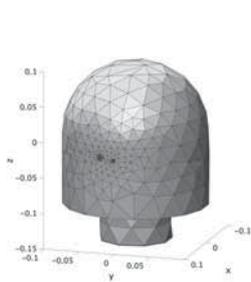


Fig.1 頭部内振動場のFEMモデル

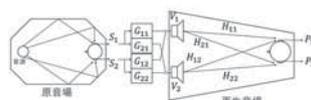


Fig.2 トランスオーラルシステム

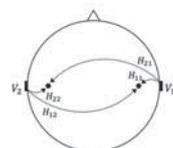


Fig.3 再生音場としての頭部内振動場

3-Q-8

3-Q-8 自動運転中の音響的マインドフルネスが ドライバに与える影響

Effects of Acoustic Mindfulness on Drivers During Autonomous Driving

☆福田千聖, 石光俊介, 大森達也(広島市大・情報科学研究科)

- ◆自動運転技術の実用化が進んでいるが、未だドライバの必要性は失われておらず、事故防止に向けた取り組みが必要である。事故の大きな原因のひとつは運転中の疲労の蓄積であり、この疲労を軽減する効果的な方法を見いだすことが重要である。
- ◆本研究では、呼吸に同期したフィードバック (FB) 音を自動運転環境下のドライバに呈示することで、音響的マインドフルネスによる疲労軽減を行うことを検討した。FB 音の有無による影響を脳波 (α 波の含有率) を用いて検証し、運転パフォーマンスを比較した。
- ◆結果として運転パフォーマンスに傾向はみられなかったものの、FB音聴取による集中力の向上が確認された。(Fig.1)

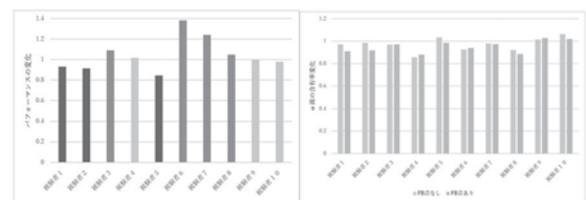


Fig.1: Driving Performance after Listening (Left) and Alpha Wave Content after Feedback Sound Listening/without listening (Right)

3-Q-10

3-Q-10 室内 3D モデルに基づく室内音響インパルス ス応答の推定とその音場制御への適用に 関する検討

Investigation into the application of room impulse responses estimated using indoor 3D models for sound field control.

☆伊藤陸人, 立蔵洋介(静岡大院・総合科学技術研)

- ◆研究概要
LIDAR スキャナで取得した室内 3D モデルを MESH2IR に入力して推定した室内インパルス応答(RIR)を用いて、マルチチャンネル音場再現を実行したいが、現状推定した RIR は実測した RIR との差が大きい。そこで、推定した RIR に後処理を加えることで実測した RIR に近づけたい。どのような処理が必要かを明らかにするために、マルチチャンネル音場再現を行う上で、RIR が保有する情報のうち重要度の高いものを検証する。検証結果をもとに、推定した RIR に施す処理の方向性を決定する。

◆実験内容・結果

1. 推定した RIR は 8000 Hz 以上の周波数成分を持たないため、8000 Hz 以下の周波数成分のみでもマルチチャンネル音場再現が可能であるか検証した。結果、制御後の信号には高周波のひずみが重畳する分 SNR が低下したものの、クロストークが抑えられており、8000 Hz 以下の周波数成分のみでも問題ないことが確認された。
2. 後部残響をカットした状態でマルチチャンネル音場再現を実行したところ、0.100 s 以降の後部残響をカットしても精度が低下しないことが確認された。そのため、推定した RIR の 0.100 s 以降の波形をカットし、ピーク位置や振幅の減衰具合を実測した RIR に近づけることを今後の方針とする。

3-Q-11

3-Q-11 深層学習による Mode Matching 法に基づいた 2.5 次元音場再現の基礎的検討

Basic study of 2.5D sound field reproduction based on mode matching using deep learning

☆川瀬敬子, 佐藤元, 津國和泉, 池田雄介(東京電機大)

◆研究背景

- Mode Matching 法 (MM 法) では, 球面調和関数の展開係数の最大次数に応じた多数のマイクロホンが必要

◆提案手法

- 深層学習を用いて少数のマイクロホンによる MM 法を提案
- マイクロホン信号から高次の球面調和関数の展開係数を推定しスピーカの駆動関数を決定

◆シミュレーション実験

- 提案手法によってより広い範囲で音場再現が可能

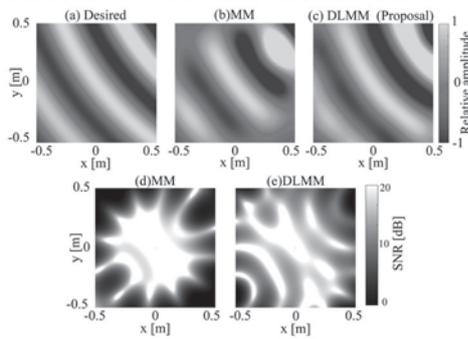


Fig.1: Reproduction sound field and distribution of SNRs (MM: N=5)

3-Q-13

講演取消

3-Q-12

3-Q-12 深層学習によるスピーカの伝達間推定を用いた Pressure Matching 法

Pressure Matching using deep learning estimation of transfer functions and original sound fields

☆堀越光樹, 佐藤元, 津國和泉, 池田雄介(東京電機大)

◆研究背景

- Pressure Matching(PM)法では多くの測定が必要とされ, 実装は容易ではない

◆提案手法

- 原音場とスピーカの伝達関数の推定を目的とした, データ駆動型深層学習による PM 法の提案

◆シミュレーション実験

- 少数のマイクロホン信号から多数の制御点における音圧もしくは伝達関数の推定
- 実験の結果, 従来の PM 法と比較して 2dB の合成精度向上

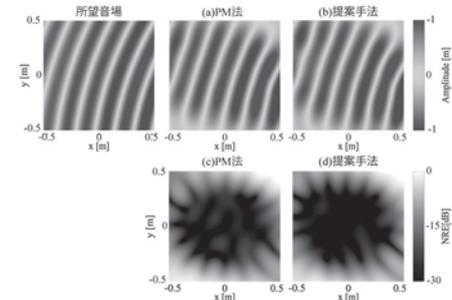


Fig.1: Sound field and normalized reproduction error (a),(c): Pressure Matching Method, (b),(d): Proposed Method

3-Q-14

3-Q-14 周波数に応じた制御範囲を用いた Pressure Matching 法に対する実測による再現精度の評価

Measurement evaluation for reproduction accuracies of pressure matching using the scalable control regions corresponding to the reproduction frequencies

☆松本綾太, 津國和泉, 池田雄介(東京電機大)

◆研究背景

- Pressure matching 法では周波数により制御可能な範囲は異なる

◆提案手法

- 周波数に応じて制御範囲を可変

◆実測実験

- 円形マイクロホンアレイと矩形スピーカアレイを用いて音場再現
- 半径 0.4 m において約 2~15 dB の精度改善

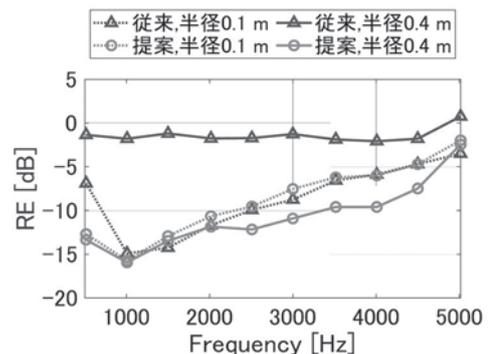


Fig.1: Reproduction error at each frequency (△: Conventional method, ○: Proposed method, Solid line: Radius 0.4 m, Dashed line: Radius 0.1 m)

3-Q-15

3-Q-15 マルチタスク深層学習による物体を含む 二次元音場画像のための 雑音除去とセグメンテーション

Denoising and segmentation for two-dimensional sound field images with occluding objects by multi-task deep learning

○谷川理佐子(NTT/早大理工), 石川憲治, 原田登(NTT), 及川靖広(早大理工)

- 光学的音響計測による二次元音場イメージング
 - 音波の伝播の様子を可視化・分析可能
 - ✓ 可視化画像にはノイズが多い
 - ✓ 物体と音波の干渉のあるデータの解析には音波と物体のセグメンテーションが必要
- 提案手法: DNN を用いた音場画像の雑音除去とセグメンテーション
 - ◆ DNN 雑音除去モデルの改良によるマルチタスク学習を実現
 - ◆ 音場シミュレーションによるデータセットの作成
 - ◆ 実験により実測データへの適用可能性を確認

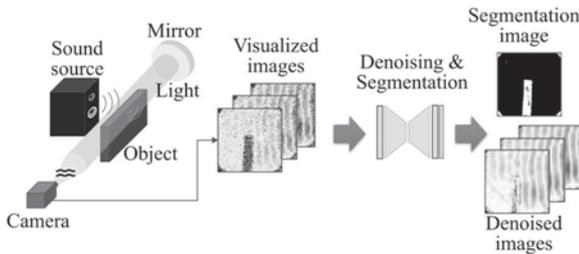


Fig. 1: Conceptual diagram of our work. Sound field with interacting object is captured as images by optical measurement system. Visualized images are converted to denoised and segmentation images by our method.

3-Q-17

3-Q-17 光学的音響計測による 散乱音場の 3 次元復元

Three-dimensional reconstruction of scattered sound field using optical measurement of sound

○石川憲治, 谷川理佐子, 原田登(NTT)

- 光学的音響計測による 2 次元音場イメージングデータから散乱音場を 3 次元復元する手法を提案
- 散乱体がない条件とある条件でそれぞれ複数方向から音場測定を行い, それらの差分から散乱音場の投影データを取得
- 球面調和関数展開を用いた外部問題として定式化
- 楕円体による散乱音場を測定し, 有限要素法によるシミュレーションとの比較を実施

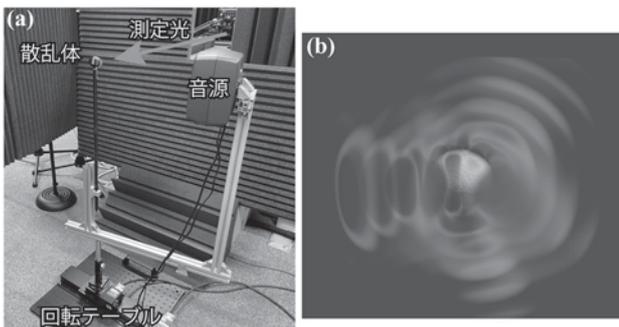


Fig.1 (a) 測定システム, (b) 20 kHz 音波入射時の散乱音場の 3 次元可視化

3-Q-16

3-Q-16 音場情報埋め込みを用いた DNN 音場画像雑音除去

DNN sound-field image denoising using sound field information embedding

☆浦田大輔, 及川靖広(早大理工), 石川憲治(NTT), 谷川理佐子(NTT/早大理工)

- ◆ 背景: 光学的手法による 2 次元音場の定量計測
 - 実測音場画像には雑音が多く含まれ, denoising が必要
- ◆ 従来手法: Deep Sound-Field Denoiser(DSFD)
 - ネットワークの入力: Noisy 音場画像
- ◆ 提案手法:
 - 入力特徴量に音場情報である波数を付加し, 波数から平面波を生成・埋め込むネットワークの機構に
- ◆ 結果: 従来手法と同等以上の雑音除去精度
- ◆ 課題: 実測音場画像の denoising の改善

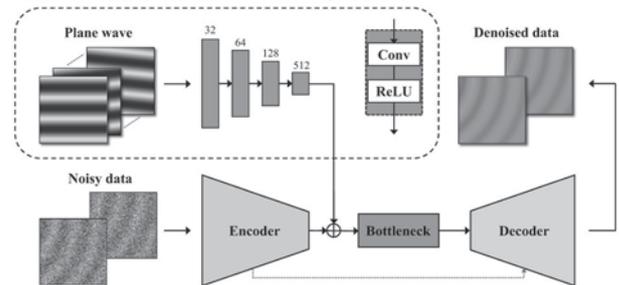


Fig 1: Outline of Proposed Network. Plane-wave images are feature-extracted by encoder within dotted line and added in first layer of bottleneck.

3-Q-18

3-Q-18 近接音場頭部伝達関数の 波動関数展開に関する一検討

A study on wave expansion of near-field head-related transfer functions

○松井健太郎(NHK)

- ◆ 相反定理に基づく頭部伝達関数 (HRTF) の波動関数展開が提案されている。HRTF の測定を外耳道入口を音源位置とする放射音場の測定とみなし, 外部問題として波動関数展開することにより, 任意位置の HRTF を外挿することが可能となる。
- ◆ 本報では, 弊所で測定した近接音場 HRTF を用いて, 実測した HRTF と外挿した HRTF との誤差を評価し, 外挿の妥当性を検証した。
- ◆ Figure 1 に実測および測定距離 0.2 m の HRTF から外挿した水平面左耳の HRTF を示す。測定距離においてはおよそ 4 dB から 6 dB の歪みで HRTF が外挿され, それより遠方でも歪みの増加は小さく抑えられることが, 測定距離より近傍では歪みの増加が明らかに大きくなることが分かった。

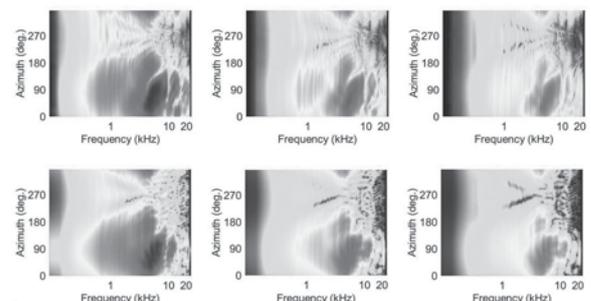


Fig. 1: Comparison of measured and extrapolated HRTFs

3-Q-19

3-Q-19 ニューラルフィールドを用いた HRIR の補間と評価

Head-related Impulse Response Interpolation Using Neural Field Method
 ☆JIANG YUNDONG, 藤平飛俊, 及川靖広(早大理工)

◆Background

The measurement of HRIR typically only covers certain angles. However, we hope to obtain HRIRs for a wider range of angles.

◆Introduction

We establish a neural field that, by training on existing HRIR data, can generate HRIRs for angles that have not been measured. We also use the generated HRIRs to render binaural audio, validating the feasibility of the interpolation.

◆Result

HRIRs generated by the neural field are close to the ground truth data. The binaural audio rendered by the interpolated HRIRs is proved to sound more continuously with more precise intervals.

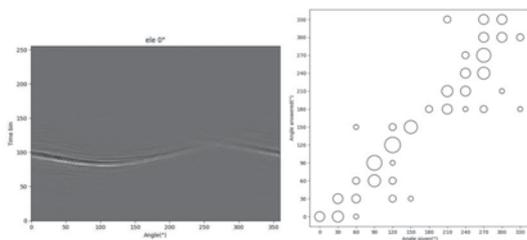


Fig.1: HRIR generated at 0° elevation angle. Fig.2: Results of the evaluation experiment.

3-Q-21

3-Q-21 音源方向の教師付け無しに学習可能な音響イベント定位

A Direction of Arrival Annotation Free Training for Sound Event Localization and Detection System

©安田昌弘, 齊藤 翔一郎, 佐藤 菜緒, 中山 彰, 原田 登 (NTT)

◆音響イベント定位 (Sound event localization and detection: SELD)

➢ 音響信号から周辺で発生した音響イベントの種類, 時刻, 音源到来方向を推定

◆課題

➢ SELD システムの学習には音源到来方向のラベルが付与されたラベル付きデータが必要
 ➢ 音源到来方向のラベリングはコストが高く, 十分な量のラベル付実データが存在しない

◆提案法

➢ 音源到来方向についてラベル付けされていないデータを用いた SELD システムの学習
 ➢ あらかじめ決められた複数の方向へのビームフォーミングにより, 各ビーム方向から到来する音が抽出された音響信号を利用
 ➢ 各方向にイベントが存在する事後確率を Multiple instance learning (MIL) の枠組みで音源到来方向のラベル無しに学習する

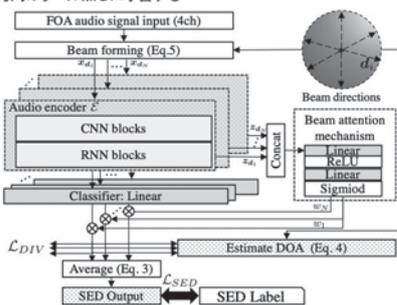


Fig. 1 Network architecture of proposed method

3-Q-20

3-Q-20 製品作動音のばらつきを考慮したロバストな異音検出手法の検討

Robust anomaly sound detection method considering variability in Product operational sound.

○坂本湧暉, 笠原太郎, △宗戸俊樹, 伊藤秀浩(アイシン)

◆自動車部品の製造現場では熟練の官能検査員による異音の合否判定が行われており, 省人化・自動化による生産性の向上が求められている. 本稿では, 検査対象となるパルス音のみを検出する方法を提案する.

◆従来では波効率(Crest factor)を用いたパルス音の検出を行ったが, 製品のばらつきによる音色の変化に対応できるよう二乗平均平方根(Root Mean Square: RMS)の変化量(Fig1)と四分位範囲を用いた異音検出手法を検討する.

◆従来手法と提案手法で見逃しと正常データの統計量 3σ を超えた割合を過検出率として比較を行った結果, 過検出数が1.24%から0.28%に減少し, 見逃しも無くなったことで精度向上を確認した

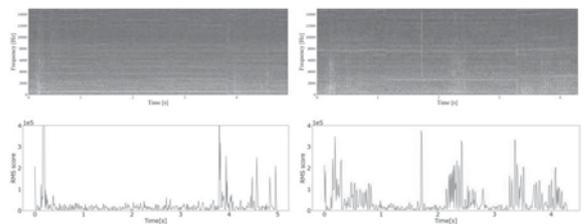


Fig.1: Spectrogram and RMS Time-Varying Envelope Curve (left: normal sound, right: anomaly sound)

3-Q-22

3-Q-22 構造を持つ雑音信号を用いた環境音分類の実データを用いない事前学習

Pre-training without real data for environmental sound classification using noise signals with structure.

○和田唯臣(長岡高校), △片岡裕雄, 坂東直昭(産総研)

◆目的: 環境音分類を少量の実音響データを用いて精度を改善.

◆従来手法: 200 万の大規模実音響データを用いた事前学習

➢ 音響認識モデル PANNs は実データ AudioSet により事前学習.
 ➢ データ収集, ラベリングに伴うコスト, 倫理的問題などが発生.
 ➢ 画像: 数式駆動教師あり学習.

✧ 教師ラベル付フラクタル画像で事前学習.

◆提案手法: Perlin 雑音を用いて音響データを自動生成・事前学習.

➢ 数式生成の教師ラベル付音響による事前学習効果を検証.
 ➢ 画像処理で良く用いられる Perlin 雑音を用いてデータ作成.

◆実験結果: Perlin 雑音の音響データ+ResNet-18 により性能検証.

➢ ESC-50 でファインチューニング → 4%の精度向上.

ResNet の表現力, 数式生成音響データと実データの乖離からマイナスの事前学習効果も一部で確認.

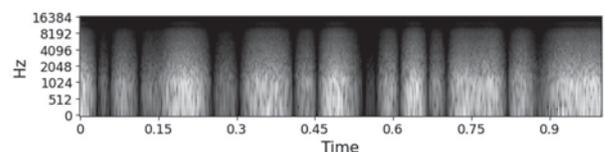


Fig.1: Example of an automatically generated noise Mel spectrogram.

3-Q-23

3-Q-23 環境音に対する 印象説明文データセットの構築と分析

Construction and analysis of impression caption dataset for environmental sound

◎岡本悠希(東大院・情報理工), 永瀬亮太郎, △岡本南美(立命館大),
齋藤佑樹(東大院・情報理工), 福森隆寛, 山下洋一(立命館大)

◆環境音聴取時に感じる印象を記述した説明文データセットを構築

- ✓ 従来の説明文: 音の発生内容や発生順序に限定
- ✓ 本データセット: 「切ない」「怖い」のような印象情報を含む
 - 3,600文の印象説明文を収録 (30音響イベント × 40音 × 3文)

◆ChatGPTを用いた印象説明文データセットの構築

- STEP1: クラウドソーシングサービスによって環境音に対する印象語を収集
STEP2: 収集した印象語を用いて ChatGPT で印象説明文を候補生成
- ChatGPTへの入力: 音響イベントラベル, 環境音に対する印象語, 極性情報
- STEP3: STEP2で生成した印象説明文候補のうち最も適切な説明文を選択
- クラウドソーシングサービスによって選択 (印象説明文の例は表1参照)

◆構築した印象説明文データセットの分析

- ✓ 環境音に対して印象説明文が妥当であるか5段階評価
 - 多くの説明文が3以上の回答を獲得 → 環境音に対して妥当な説明文であると確認

表1: 印象説明文の例

音響イベント	ChatGPTへの入力			出力された説明文
	快 or 不快	落ち着いた or 緊張した	印象語	
小鳥のさえずり	快	落ち着いた	鋭い	鋭く澄んだ音が耳に届き、心に穏やかな安らぎをもたらす印象
教会の鐘の音	快	緊張した	華やかな	華やかな音色が鳴り響き、新たな始まりを予感させる喜びの音
木のドアを叩く音	不快	落ち着いた	怖い	怖さを感じさせるその重い響きは、心に緊張と不安をもたらす
掃除機の音	不快	緊張した	うるさい	うるさく響き、一瞬で心をざわつかせる雑音

3-Q-25

3-Q-25 学習済み環境音認識モデル に対する学習レス適応

Learning-less adaptation of pretrained models for environmental sound classification

◎砺波 紀之(NEC Corporation), 井本 桂右(同志社大学),
美島 咲子, △荒井 友督, 近藤 玲史, △榎野 智之(NEC Corporation)

- ◆学習済み環境音認識モデルに対して、最適化を用いない適応手法を提案する。
- ◆従来の適応手法(例: fine tuning や転移学習)では、Graphics Processing Unit (GPU) を用いることが前提条件であった。
- ◆提案手法では、深層学習モデル中間層に発現する時間周波数域的構造に着目し、中間層において周波数フィルタリングを模擬することで、最適化を必要としない少演算量な適応手法を提案する。
- ◆ECS-50を用いたイベント分類性能結果より、提案手法を用いることで、最大で20.40%ポイントの分類精度向上を確認した。
- ◆また、図に示すように、モデル中間層において時間周波数域的表現が発現されることを定性定期に確認した。

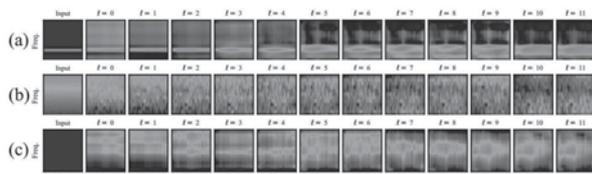


Fig. 1: time-frequency-ish representation of intermediate layers for (a) pure tone, (b) white noise, (c) silent signal

3-Q-24

3-Q-24 埋め込み共有による分散協調異常音検知

Distributed collaborative anomalous sound detection by embedding sharing
◎土肥宏太, 川口洋平(日立製作所)

- ◆機械の稼働音から自動点検を行う異常音検知技術を開発している。産業機械の異常音検知では、各クライアントが異なる機械や異なる運転状態のデータを有する。そのため、各クライアントの生データを秘匿しつつ複数クライアントのデータを活用する際に、連合学習等の既存手法の適用が困難である。
- ◆提案方法では、各クライアント共通の事前学習モデルを使用して生データの埋め込みを計算し、各クライアントにおいて計算された埋め込みをサーバーで集約した上で、外れ値暴露を用いて異常音検知モデルを学習する。これにより、各クライアントによる生データの共有を行わずに、複数のクライアントのデータを活用しながら異常音検知モデルを学習することが可能になる。
- ◆実験において、提案手法は、埋め込みを共有しないベースラインに比べて、異常音検知精度(AUC)を平均で6.8%向上させることが分かった。

Table 1: Anomaly detection results for different machine types. "total" denotes the average AUC for all 23 machine IDs. (OpenL3-512) and (OpenL3-6144) denote OpenL3 with embedding dimension of 512 and 6144, respectively.

Model	VAE	VIDNN	Glow	Proposed (YAMNet)	Proposed (PANNs)	Proposed (OpenL3-512)	Proposed (OpenL3-6144)
ToyCar	83.3	81.2	75.9	78.9	79.0	87.5	91.4
ToyConveyor	74.1	74.7	70.7	63.4	63.1	67.8	70.2
fan	72.1	72.8	71.5	68.6	67.3	70.5	76.4
pump	76.2	76.6	76.3	71.1	76.0	81.6	82.4
slider	84.2	84.7	93.9	82.2	78.8	96.2	95.6
valve	70.2	83.0	86.1	89.4	89.6	94.3	97.3
total	76.8	79.0	79.4	76.1	76.2	83.7	86.2

3-Q-26

3-Q-26 CNNを用いた音響による接近車両検出に おける音データの前処理方法の検討

A Study on Methods of Preprocessing Audio Data for Approaching Vehicle Detection Using Sound Based on CNN

☆伊藤隆佑, 堀江拓矢, 加藤朋暉, 旭健作, 坂野秀樹(名城大院)

- ◆出会い頭衝突事故の防止を目的に、道路上で収録した音を用いた接近車両検出を検討している。本研究で重要となる車の接近音は低周波帯域に多く含まれる。そのため、その帯域を捉えやすくするために音データの前処理方法に関して検討を行い、接近の検出精度向上を目指す。
- ◆音データの標準化周波数を従来の48[kHz]より低くしたところ、18[kHz]付近で全体的な検出精度が最も向上した。
- ◆更に、データの前処理として作成するメルスペクトログラムのメルバンド数を従来の32より大きくし、低周波成分の分解能を向上させることを試みた。結果として、メルバンド数とともにCNNの畳み込み層におけるkernelサイズを大きくすることで精度が向上し、メルバンド数64, kernelサイズ(6×6),(14×14)にすることで、最接近時刻(0.0秒)の4.5秒前から85%以上の精度で接近車両を検出可能となった。

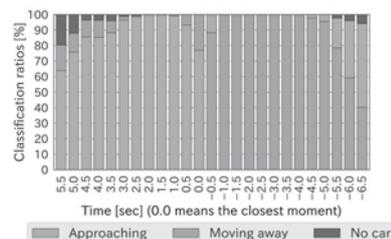


Fig. 1: Classification proportion at each time in approaching and moving away situations

3-Q-27

3-Q-27 音による交通モニタリングの性能向上の検討 —DCASE Challenge Task 10 への参加を通して—

Performance Improvement in Acoustic-Based Traffic Monitoring:
Through Participation in DCASE Challenge Task 10

☆高橋知宏(都立大), 植野夏樹(熊本大/都立大),

木下裕磨(東海大), 若林佑幸(豊橋技科大), 小野順貴(都立大)

- ◆DCASE 2024 Challenge Task 10: Acoustic-Based Traffic Monitoring
 - 車種と進行方向ごとの通過車両数を推定するタスク
 - 評価指標として Kendall's Tau Rank Correlation (KTRC) を使用
- ◆KTRC を直接最適化するための損失関数は知られていなかった
- ◆KTRC の精度向上を目的とした新たな損失関数を提案
 - マッチング損失の概念を順位関数に適用することで導出
 - データの順序関係の一致度を真値と推定値の間で高める
 - ノンパラメトリックな尺度であり様々な応用が期待できる
- ◆提案した損失関数による推定誤差の改善を確認 (Fig. 1 参照)

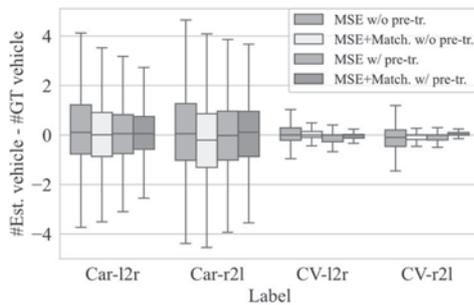


Fig. 1: Error with and without Matching and pre-training for location 6

3-Q-30

First-shot 異常音検知のための評価用データ ToyADMOS2#の収録

ToyADMOS2# (sharp) Dataset for the First-Shot Anomalous Sound Detection

○原田登, 仁泉大輔, 大石康徳, 竹内大起, 安田昌弘(NTT)

- ◆First-shot 異常音検知タスクの概要と, 新規に収録したデータセット ToyADMOS2#, およびベースライン実装による評価結果を紹介する。
- ◆(a)おもちゃの自動車 (ToyCircuit), (b)小型ドローン(HobberingDrone), (c)ドライヤー(HairDryer), (d)電動歯ブラシ (ToothBrush)の4種類の機器の稼働音を新たに収録した。

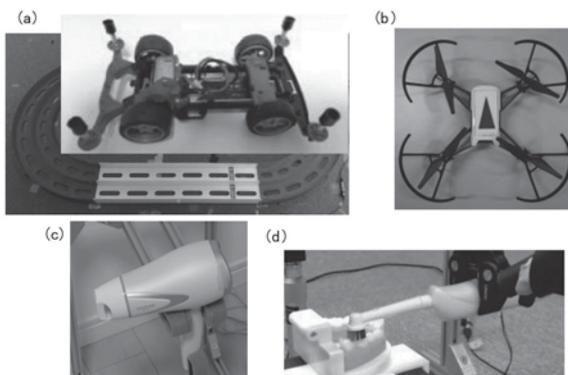


Fig.1: Machine models.

3-Q-29

3-Q-29 音響シーン・同期・マイク位置の手がかりのための多チャンネル録音を用いたアンカー音検出の検討

Anchor sound detection using multi-channel recordings for acoustic scene, synchronization, and microphone location cues

○河村隆生, 小野順貴(都立大)

- ◆背景: 分散録音で空間情報を用いるために同期・自己位置推定が必要
 - 校正信号のような既知の音の利用は効果的だが事前に準備が必要
- ◆着眼点: 身の回りに校正信号となりうる音があるのではないかと?
 - 常に同じ波形で同じ位置から発せられる音(例えばテレビの起動音)をアンカー音と定義し, 音響シーン・同期・マイク位置の手がかりとなることを期待
- ◆提案: 「波形の類似性」「時間差の一致度」を考慮したアンカー音検出

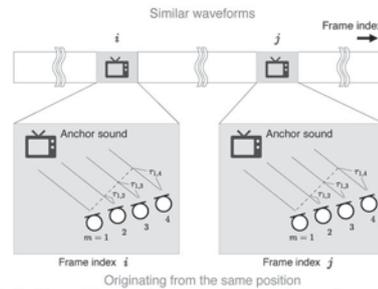


Fig.1: Definition of the anchor sounds: They have the same/similar waveform and originate from the same position.

- ◆実験: 実環境データセット SINS database へ提案法を適用
 - テレビの起動音やドアを閉める音が検出されることを確認
 - マイクロホンアレイの向き推定への応用可能性を検討

3-Q-32

3-Q-32 Long Short Term Memory Autoencoder(LSTMAE)を用いた豚呼吸器病の早期異常検知

Early detection of abnormalities in swine respiratory disease using Long Short Term Memory Autoencoder (LSTMAE)

☆の場悠基, 石光俊介, △湯本誠司(広島市立大院), △三上修,

△小川洋介, △井上寛暁, △石田三佳(農研機構), △田中大介(新居浜高専)

- ◆近年, 日本の畜産業は拡大している。この状況下において, 家畜の伝染病発生は重大な脅威であり, 大きな被害をもたらす可能性がある。また, 肺炎などの呼吸器病罹患は成長を妨げ, 農家の費用対効果を下げている。よって, 罹患を早期発見するシステムが必要である。
- ◆本研究では, 体内伝導音のスペクトログラムを入力とした Long Short Term Memory Autoencoder (LSTM オートエンコーダ)を用いて, 個体専用の機械学習モデルを構築し, 個別に状態を検知した。
- ◆動物実験における豚3頭を対象にそれぞれ機械学習モデルを構築し, 健常と罹患を識別した。そのうちの2頭の識別結果を以下に示す。全体的に約80%以上の精度を識別でき, pid5 から pid7 にかけて約80~90%以上の精度で罹患を識別できた。

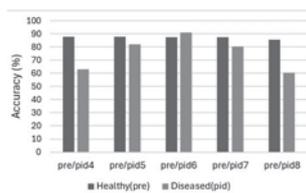


Fig.1(a): PigA identification results

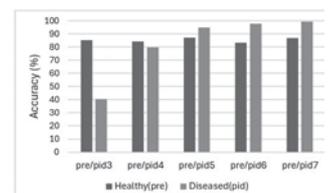


Fig.1(b): PigB identification results

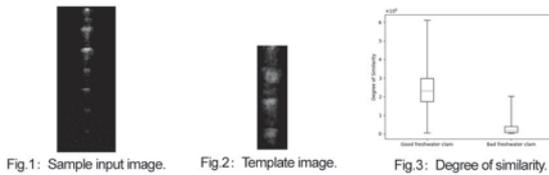
3-Q-33

3-Q-33 超音波画像を用いたテンプレートマッチングによるしじみの良否判別の試み

Attempt to discriminate good or bad freshwater clams by template matching using ultrasonic images

☆古川尚幸(近畿大院), 篠原寿広, 上保徹志, 中迫昇(近畿大)

- ◆しじみは出荷前に漁師の手作業によって出荷可能な貝(よしじみ)か、出荷不可能な貝(不良しじみ)かに判別されている。しかし、この作業は集中力を必要とする長時間の労働である。
- ◆本研究では、超音波画像を用いてしじみの良否を高精度に判別する可能性について、画像処理やテンプレートマッチング、統計処理の観点から検討を行う。
- ◆Fig.1 はよしじみの超音波画像をトリミング処理した入力画像の例、Fig.2 はよしじみの超音波画像をトリミング処理したテンプレート画像である。本研究では、よしじみ・不良しじみの入力画像に対し、テンプレート画像との類似度が最大となる値を入力画像ごとにそれぞれ調べた(Fig.3を参照)。
- ◆Fig.3 よりよしじみの、テンプレート画像との類似度の値が不良しじみより大きい傾向があることから、しじみの良否判別の可能性が示唆された。



3-Q-35

3-Q-35 光レーザマイクロホンを用いたマルチチャネルフィードフォワードアクティブノイズコントロールシステムの実環境下における騒音低減性能の実機評価

Experimental Evaluation of Noise Reduction Performance of Multichannel Feedforward Active Noise Control System with Optical Laser Microphone in Real Environment

☆水谷真絃(立命館大院), 岩居健太, 西浦敬信(立命館大), 添田喜治(産総研)

フィードフォワードアクティブノイズコントロール(FFANC)システムにおける因果律の制約を緩和するために、光レーザマイクロホンを用いた FFANC システムが提案されている。しかし、取得された参照信号は、空間の反射特性の影響を受けにくいいため、残響成分を十分に含まない。そのため、残響環境では、この ANC システムの騒音低減性能が低下すると考えられる。本稿では、空間の反射特性を考慮するために、光レーザマイクロホンと気伝導マイクロホンを用いたマルチチャネル FFANC システム(Fig. 1)を提案する。実環境での実機実験の結果、提案システムは従来の ANC システムと同程度の騒音低減性能を持つことが示された (Fig. 2)。

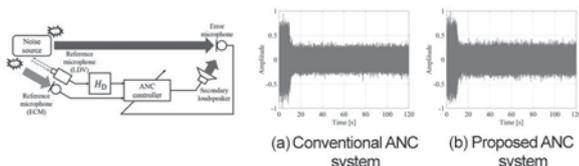


Fig. 1: Structure of proposed ANC system Fig. 2: Time waveforms of error signals

3-Q-34

3-Q-34 2つの観察点の信号を1つの誤差信号とした適応振動制御

Active window vibration control using single error signal added the vibration signals at two observation points.

☆中島尚哉, 武藤憲司(芝工大)

- ◆我々のこれまでの研究では、二次経路の長さが等しいときの2か所の観察点の振動信号を加算して1つの誤差信号として制御した。
- ◆本研究は窓振動低減システムのために、2か所の観察点の二次経路に差をつけた場合の制御効果を検討した。振動観察点A,Bを設定したうえで、単体の結果と加算した場合の結果を比較した。
- ◆音圧レベル低減量は2か所の信号を加算した時、窓全体の平均1.5 dBが最も良い結果となった。音圧レベル測定点40地点のうち、31地点で小さくなった。これより、2か所の観察点の二次経路に差をつけた場合でも窓振動低減に効果があるとわかった。

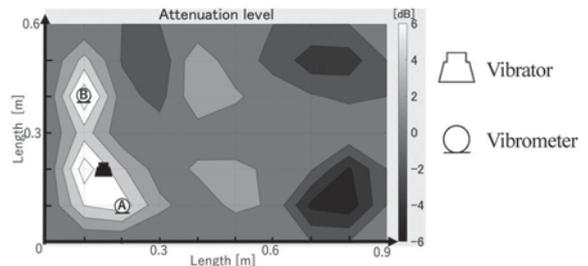


Fig. 1: Distribution of attenuation level of sound pressure near window surface.(A+B)

3-Q-36

3-Q-36 頭部反射を考慮したテトラ型マイクアレイによる音圧外挿ANC

Active noise control system with tetrahedral microphone array based on sound field extrapolation considering head reflection

☆吉松亨真(電通大), 伊藤弘章,

小塚詩穂里, 鎌土良良(NTT), 羽田陽一(電通大)

- ◆耳を塞がないANCシステムの実現に向けて、テトラ型マイクアレイを耳位置近傍に配置し、球面調和展開を用いて耳位置の音圧を予測する手法を提案する。
- ◆2つめのテトラ型マイクアレイを用いて予測誤差を補正するフィルタを逐次推定し、予測性能を向上させる手法も提案する。
- ◆テトラ型マイクアレイは騒音信号が水平面以外から到来する場合も予測可能であり、マイクアレイの外接球の表面に予測位置を配置することで実数計算による音圧予測も可能である。
- ◆様々な騒音到来方向における騒音低減性能を計算機シミュレーションにより確認した。

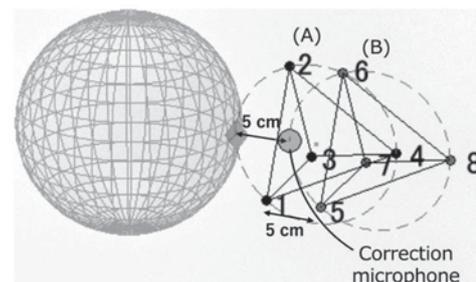


Fig. 1: Setup of two tetrahedral microphone array

3-Q-42

3-Q-42 イヤーセンタリングが近距離頭部伝達関数の合成精度に及ぼす影響

The Effect of Ear-centering on the Synthesis Accuracy of Near-Field Head-Related Transfer Functions.

☆北村航太, 坂本修一 (東北大通研/院情科研)

- ◆本研究ではイヤーセンタリングによって生ずる幾何的な問題に対処するべく、イヤーセンタリングを適用した距離変換フィルタ (DVF) による近距離 HRTF 合成法に角度補正を加え、その合成精度の評価を行った。
- ◆角度補正では、測定された HRTF のサンプリング点について、頭部中心からの見込み角と耳からの見込み角を対応させ、原点位置に合わせた角度にサンプリング点がラベリングされるようにした。
- ◆シミュレーションから求めた HRTF を目標値として合成精度比較を行った結果、角度補正を行った場合は行わなかった場合に比べ、合成精度が向上した。また、低域では通常の DVF を適用した場合よりも合成精度が向上した。

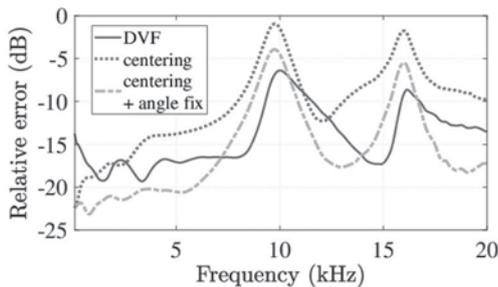


Fig.1: RMSE between synthesized HRTFs and original HRTFs

3-Q-44

3-Q-44 グラフ信号処理を用いた音源放射特性の内挿

Interpolation of source radiation characteristics using graph signal processing

◎大久保翔太, 堀内俊治, △渡邊良亮 (KDDI 総合研)

- ◆音源放射特性において、限られた位置で収録された音響信号から、収録していない位置の音響信号を補間することを目的とし、グラフ信号処理を用いた補間手法を導入した。
- ◆音源中心の水平方向に対して、音響特徴量のうち音圧レベル差、到達時間差、LSD を用いて定量評価を行った。
- ◆音圧レベル差 (Fig. 1) から、線形補間、円調和関数と比較し、90° から 270° にて精度が向上されることが確認された。
- ◆到達時間差 (Fig. 2) の結果から、円調和関数と比較し線形補間、グラフ信号処理ベースの補間精度が高いことが確認された。

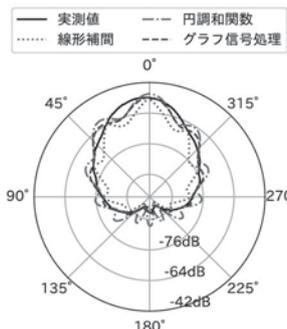


Fig.1: RMS values of captured and interpolated signals

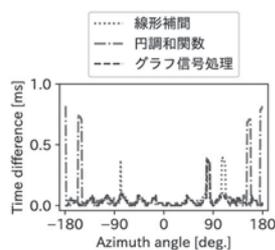


Fig.2: Time difference between captured and interpolated signals

3-Q-43

3-Q-43 発話放射特性と種々のスピーカ放射特性との比較

Comparison of speech directivity and various loudspeaker directivities

☆横田康介, 羽田陽一 (電通大)

- ◆実際に人が話しているような臨場感を再現するためには、3次元発話放射特性を再現することが重要である。
- ◆本報告では、実測に基づき音韻を平均化した実話者の発話放射特性と、複数の形状の異なるスピーカの放射特性を比較する。
- ◆Fig.1は実話者、Fig.2は球スピーカ、Fig.3は前円筒後半球スピーカ、Fig.4は実話者の頭部を3Dスキャンした疑似頭スピーカである。
- ◆水平面・正中面における放射特性の比較、球面展開係数により放射特性に含まれる各実球面調和関数の大きさの比較を行った。
- ◆比較の結果、いずれのスピーカも放射の強さが異なる方向があったが、各スピーカに胴をつけることによって放射特性がより似ることがわかった。
- ◆最も実話者の発話放射特性と類似していたのは疑似頭スピーカの放射特性だったが、単純化された形状でも部分的に実話者と似た傾向があることがわかった。



Fig.1: Speaker Fig.2: Sphere loudspeaker Fig.3: Front cylinder back hemisphere loudspeaker Fig.4: Dummy head loudspeaker

3-Q-45

3-Q-45 残差接続と物理法則を用いた深層学習による室内インパルス応答推定

Room impulse response estimation by physics-informed neural networks using residual connection.

☆倉田駿, 佐藤元, 津國和泉, 池田雄介 (東京電機大)

- ◆背景
 - ・近年、深層学習と物理法則を用いた少数のマイクロホンから初期室内インパルス応答(RIR)の推定が行われている。
 - ・2種類の損失関数を用いるため2つの損失関数の勾配が必ずしも同方向にならず学習がうまく進まない場合があったが Dynamic Pulling Method を利用することで解決してきた。
 - ・より高い精度の初期 RIR の推定を行うために、物理法則を用いた深層学習モデルに残差接続を導入する。
- ◆シミュレーション実験

測定点 144 個から推定点 400 個を提案手法と従来手法を用いて推定 ⇒提案手法は DPM の手法と比べ NSME 1.3dB 改善

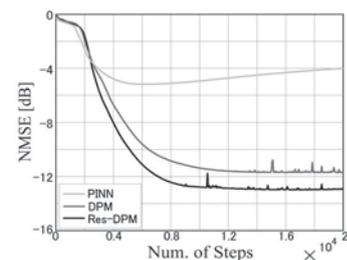


Fig.1: Number of training steps and NMSE for the proposed (Res-DPM) and conventional (PINN, PINN using DPM) methods

3-Q-46

3-Q-46 位相干渉を用いた音響測距に基づく 位置・概形イメージング～マイコンを用いた システムの小型化の検討～

Sound field imaging to identify a target position and approximate shape
using acoustic distance measurement based on phase interference.
～Consideration of compact system using microcontroller.～

☆田邊広大, 上保徹志 (近畿大), 中山雅人 (阪産大), 中迫昇 (近畿大)

- ◆本稿では、位相干渉に基づく音響測距 (Acoustic Distance Measurement : ADM)法に着目する。シミュレーションを通じて、システムの小型化の検討を行う。
- ◆1chの音響測距を用いたイメージング法に対して、シミュレーションを行い、システムの小型化を検証する。スピーカーを1つ (位置は固定)、マイク素子数を4、マイク素子間隔を0.33mとして、それぞれのマイク素子位置で位相スペクトルと振幅スペクトルを得る。これらの結果を用いてイメージングを行い、対象物の位置の推定を行う。
- ◆Fig.1はそれぞれの4つのマイクロホン素子位置で得られる距離に関する位相スペクトルと振幅スペクトルを、Fig.2はこれらの位相情報と振幅情報を用いてイメージングを行った結果を示す。マイクが4個あるため4個の円状のイメージが出力され、重なる部分が明るくなり、対象物の位置を表していることから、スピーカーの位置を固定した場合の検証ができた。

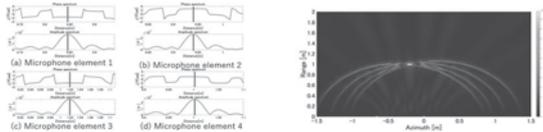


Fig.1 : Amplitude spectra and phase spectra.

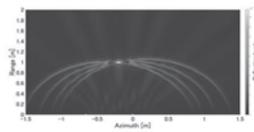


Fig.2 : Result of imaging simulation.