

1-1-1

1-1-1 トライアングルの一辺の長さ・太さとその打音・余韻の関係

Relationship between the length and thickness of one side of a triangle and its prompt and after sounds

○向日勲, △太田悠翔, △林口晴陽, △平川仁太郎(瑞陵高校理数科), 鮫島俊哉(九大・芸工), △山崎辰雄(瑞陵高校)

- ◆現在のトライアングル制作は制作者の経験によりその形状が決定されており、形状決定に関する研究がない。
- ◆トライアングルはピーターを変えると音が変わることが経験上知られている。これを衝突現象によるものであると推測し、その音の特徴は打音に顕著に表れると推察された。
- ◆打音から余韻へかけて減衰する音に注目して解析を行った。
- ◆金属棒についての解析では金属棒が細くなるほど打音から余韻へかけて減衰する音が多くなる傾向がみられた。これはたわみやすに起因するものと考えられる。
- ◆トライアングル (Fig. 1) についての解析では金属棒ほど優位な差は見られなかった。これは支持方法によるものであると考えられる。
- ◆これらのことから打音が余韻へ一定の影響を与えているものと考えられる。



Fig.1: The triangles produced for research.

1-1-3

1-1-3 スペクトル分布の時間変化分析に基づく打楽器演奏の音質評価

Sound Quality Evaluation in Percussion Playing by the Temporal Changes in Spectral Spread

☆鳥居広, 及川靖広(早大理工)

- ◆背景
打楽器演奏初心者にとって、自らの感覚で試行錯誤を重ね、演奏音の音質を改善させたり、変化させたりする作業は難しい。
- ◆概要
経験によらず、音質改善ができるシステムを目的とする。スペクトル分布の時間変化分析により演奏音をリアルタイムで評価し、練習に活用できるかを検証した。
- ◆結果
練習において数値が改善される様子が確認できた。アンケートで練習の助けになったとの回答が得られた。→練習における有効性が示された。

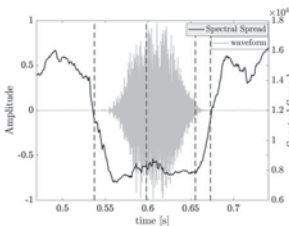


Fig.1: Example of detecting the time-points

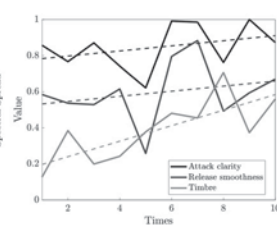


Fig.2: Changes in each feature for each playing

1-1-2

1-1-2 音叉のヤング率が振動継続時間に与える影響の検討

Effects of Young's modulus on vibration duration of tuning forks

☆上野嵩旺, 大塚 翔, 中川誠司 (千葉大)

- ◆音叉の使用用途を鑑みると、基音はより長く持続し、倍音はより速やかに減衰することが望ましい。しかし、音叉の製作工程では持続時間の評価は行われておらず、音叉振動の持続時間には個体差が存在する。
- ◆これまでに、音叉振動の持続時間に対する保持強度、温度、打撃位置等の影響が調査されてきたが、振動の持続時間への影響が予想される音叉の形状やヤング率などの材料特性といったパラメータについては検討されていない。
- ◆本研究では、音叉振動が片持梁モデルで記述できると仮定した上で、複数種の金属材料製の音叉を打撃した際に発生する音と音叉の寸法からヤング率を算出し、持続時間との関係について検討を行った。
- ◆220 Hz のステンレス製音叉を除き、ヤング率が大きいほど振動継続時間が短くなる傾向が見られた (Fig.1)。この結果は、音叉製造後の音叉のヤング率を計測することで、振動の持続時間を大まかに評価できる可能性を示唆している。
- ◆一方、220 Hz のステンレス製音叉では、ヤング率は小さいが持続時間も比較的短いという、他の音叉とは相反する結果が得られた。この原因の解明にはさらなる検討が必要である。

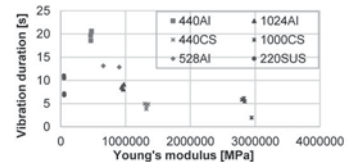


Fig.1 Young's modulus vs. vibration duration (AI: Aluminum, SUS: Steel Use Stainless, CS: Carbon steel)

1-1-4

1-1-4 Measurement of Hand Motion in Kotsuzumi Performance using Smartphone with Video Analysis

Saha Chaity, Tomohisa Funakura, Akihiro Matsutani, Marie Tabaru (Tokyo Institute of Technology)

- ◆The Kotsuzumi is a traditional Japanese drum where the structure influences sound variation, but the playing technique and striking motion are also essential. This research focuses on studying hand motion and playing techniques through an automated process.
- ◆This research was conducted without using any sensors, utilizing a smart phone (iPhone 15 Pro Max, Apple), computer vision technology, and video analysis (Fig.1), thereby not interfering with the performer's performance.
- ◆At the moment of moving the right hand to an upward direction, one hit is divided into two different stages and observed arm velocity (V_a), "arm+hand" velocity (V_{ah}), and hand velocity (V_h) (Table 1) to understand the relationship between striking techniques and sound production.

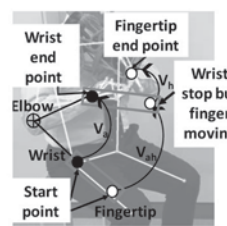


Fig. 1 Automatic detection of hand different points using computer vision technology.

Table 1 Velocity measurement for V_a , V_{ah} and V_h for different sounds.

Average Velocity	Chi (cm/s)	Pu (cm/s)	Ta (cm/s)	Po (cm/s)
arm (V_a)	50	54	55	58
"arm+hand" (V_{ah})	64	73	76	80
hand (V_h)	175	194	244	275

1-1-5

1-1-5 膜鳴楽器の打面振動モードの基本的特徴を踏まえた奏法メソッド

Methods of playing membrane instruments based on the basic characteristics of their percussion surface vibration modes

○深町浩司(愛知県立芸大・音楽学部)

- ◆子ども達の吹奏楽演奏などで、トムトムのハイスピードストローク技術だけに関心が集中し、トムトムの音色に対する関心が薄れているという懸念がある。
- ◆トムトムの音色に対する関心を高めるために、トムトムの打点と打面振動モードと音色を結び付けた。
- ◆自動打撃装置によってトムトムの打点4箇所の音響解析を行った。
- ◆新メソッドを開発して概要を示した (Fig.1)。
- ◆新メソッドを用いて、シンプルなリズムと、実際の楽曲の演奏を行い、リスニングアンケートと音響解析によって新メソッドの効果を検証した。

Fig.1 New Method Overview for tom-tom

Desired Effects	Volume...	louder	↔	softer
	The length of the sound...	shorter	↔	longer
	Attack...	sharpen	↔	soften
	Tone...	darken	↔	Brighten it up.
beating point that produce the desired effect		center of surface	↔	edge side of surface
vibration mode change		simplify (0,1)	↔	complications (1,1) etc.

1-1-7

1-1-7

高速度カメラによるエレキギター弦振動の測定とピックアップ出力の推定

Measurement of electric guitar string vibration by high-speed camera and estimation of pickup output

☆新美智也, 安藤元暉, 山田宏樹, 矢田部浩平(農工大)

- デフォーカスを利用した、高速度カメラによるエレキギター弦の二次元振動の測定手法を提案
- F値を小さくすることで、計測点が強調され振動をトラッキングしやすくなる。(Fig.1)
- 測定データから出力信号をシミュレーション
- シミュレーション信号と録音音源を比較し提案手法の妥当性を検証



Fig. 1 Guitar string images with different F-numbers .

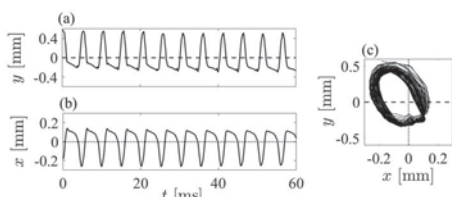


Fig. 2 Vibration of the plucked guitar string displacement: (a) the vertical displacement, (b) the horizontal displacement, and (c) the x-y plane.

1-1-6

1-1-6 ソリッド・ギターモデルにおける撥弦の減衰特性に関する研究

Experimental study on the damping characteristics of strings in a solid guitar model

○稲本絢子(京工繊大院), 田中洋介(京工繊大)

- ◆弦楽器において、デッドポイントという減衰の早い点が存在する。このような減衰の違いが発生する原因として、ネックの振動が弦の振動を吸収していることが挙げられる。ソリッドギターにおいては、音が長く一緒に続くことが品質の一つとされているため、デッドポイントの存在は問題である。本研究では弦楽器の中でもエレキベースを対象とし、簡易モデルを用いて弦とネックの振動の計測システムを構築し、減衰特性を時定数として測定した。
- ◆本実験で構築した計測システムを Fig.1(a)に示し、5フレットを押弦した弦とネックの時間変位を Fig.1(b)に示す。得られた時定数はそれぞれ、0.896 sec と 1.420 sec であった。これにより、ソリッド・ギターモデルにおける減衰特性を定量的に評価することが可能になった。

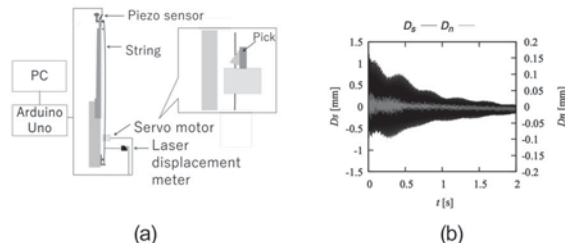


Fig.1 Measurement of displacement of string and neck: (a) Experimental system, and (b) Time variance of string and neck displacement

1-1-8

1-1-8 ボディ内の音場を考慮したアコースティックギターの有限要素法による音響振動連成解析

Vibro-acoustic analysis of an acoustic guitar involving the sound field in the body using the finite element method.

☆長谷良樹, 鯨島俊哉(九大・芸工)

- ◆ギターのFEM振動解析と、ボディ内音場の1自由度モデルを組み合わせた音響振動連成解析を行った。提案する物理モデルを Fig.1 に示す。
- ◆(a)ボディ内音場との連成を含むモデル、(b)ボディ内音場との連成を含まないモデル、(c)ボディ内音場との連成を含むサウンドホールを塞いだモデルについて、弦を含まない状態でのインパルス応答を調べた。ブリッジのz軸方向振動速度の時間応答を Fig.2 に、周波数応答を Fig.3 に示す。
- ◆弦を含んだ解析も行ったが、高域での解析精度などに課題が残った。

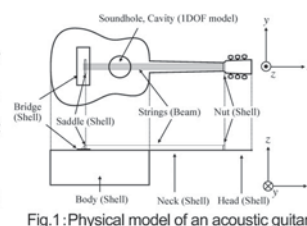


Fig.1:Physical model of an acoustic guitar.

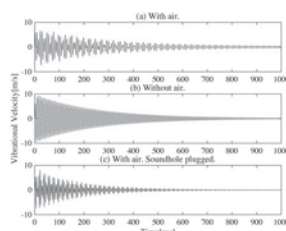


Fig.2:Time history of vibrational velocity.

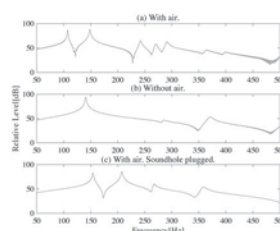


Fig.3:Frequency response of vibrational velocity.

1-1-9

1-1-9 三線の心が音と共鳴胴の膜振動に与える影響の評価

Evaluation of the effect of shin of a sanshin on the sound and the membrane vibration of the sound box.

○小島知也, 西宮康治朗, △長秀雄 (青山学院大)

- ◆三線は沖縄の伝統ある民族楽器の1つである。
- ◆心という部位が三線の音色にどのような影響を与えているか評価することを目的として、三線胴の打撃試験により、振幅の減衰波形や膜の振動のモード形状から心による音と膜振動の変化を評価した。
- ◆1000 Hz以下の周波数では減衰やモード形状に変化は見られなかったが、それより高い周波数になると、心の挿入によって音や膜振動の減衰率が小さくなった。
- ◆Fig. 1に1740 Hzのモード形状について示す。今回使用した膜は左半分が振動しにくい状態だったが、心ありの場合モード形状はFig. 1の右下に示すような理想的なモードに近づくことが分かった。

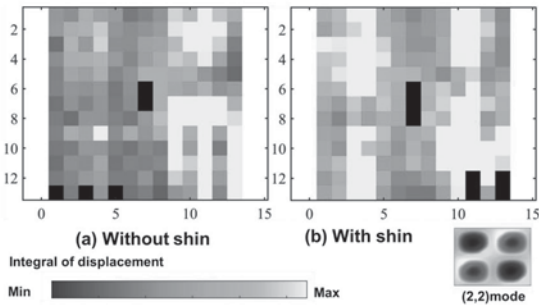


Fig.1: Membrane vibration mode on 1740 Hz.

講演取消

1-1-11

1-1-11 圧電素子によるバイオリン表裏板振動の非破壊分析と検証

Non-destructive analysis and verification of the vibration of the front and back plates of a violin using piezoelectric elements

○近藤大智(明和高), △加賀春樹, 中村謙之(明和高), 亀川徹(東京芸大)

- ◆クラドニ図形と圧電素子を使用してバイオリンを分解したり傷つけたりすることなくバイオリンの板の振動を解析した。
- ◆圧電素子とマイクロホンを使用し、バイオリンの板の振動への圧電素子の影響を考察した。
- ◆肩当てを装着しない場合とした場合の違いを、圧電素子による測定結果を用いて比較した (Fig.1)。
- ◆肩当ての有無の違いを聴取実験によって比較したところ、11名中9名が肩当ての有無を正答し、6名が肩当て有り心地良いと感じたという結果が得られた。

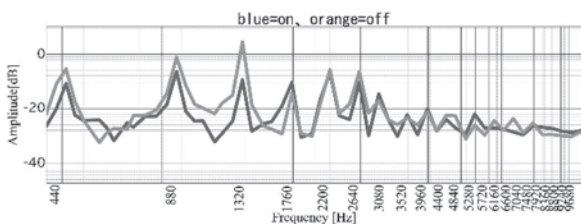


Fig.1: microphone data when wearing shoulder rest or not after Fourier transform without piezoelectric element

1-1-10

1-1-12 エンドピンストップパーの形態によるチェロの擦弦振動への影響

Effect of endpin stopper form on cello bowed string vibration

☆上條翔太郎, 鮫島俊哉(九大・芸工)

- ◆エンドピンストップパー(Fig.1)を機械的支持条件(Fig.2)として模擬
- ◆チェロを梁振動場またはシェル振動場として物理モデル化(Fig3)
- ◆等価回路表現による擦弦振動場と有限要素法によるチェロの振動場を連成し、擦弦点の振動変位の時間波形(Fig.4)を計算
- ◆支持バネの剛性の大きさがヘルムホルツ波の形成時間に影響

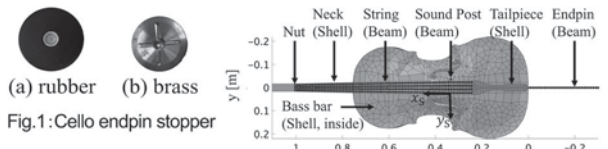


Fig.1: Cello endpin stopper

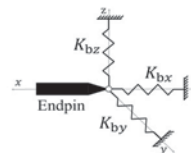


Fig.2: Boundary condition of endpin tip

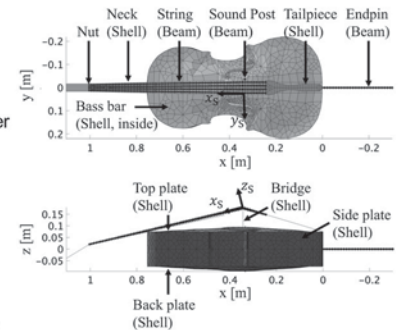


Fig.3: Physical model of cello

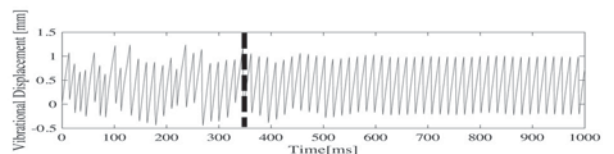


Fig.4: Vibrational displacement of the excitation point

1-1-13

1-1-13 撥弦楽器における ピック-弦-ブリッジ連成系の振動解析

Vibration analysis of pick-string-bridge coupled system
on plucked string instruments

◎塩屋仁菜(九大・芸工), 鮫島俊哉(九大・芸工)

- ◆著者らは、弦とピックによる撥弦物理モデルとその解析方法を提案している。
- ◆本研究では、弦を張力を有する梁振動場、ピックをシェル振動場、ブリッジを3次元弾性体として物理モデル化し、FEMによる離散化を行い、それらの相互作用をすべて考慮して適切に連成させた、新たな撥弦物理モデルを構築した。
- ◆撥弦物理モデルを用いて計算し、従来の撥弦モデルと比較した。
- ◆また、実物の箏を人間が撥弦する際のピックの動きを測定し、それに基づいたピックの運動軌跡を計算条件に反映することを試みた。

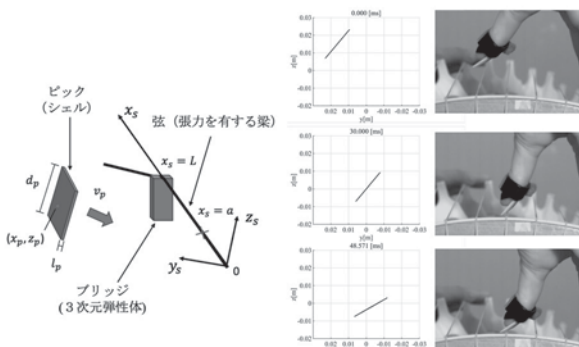


Fig.1: Geometrical arrangement of the pick, string and bridge.

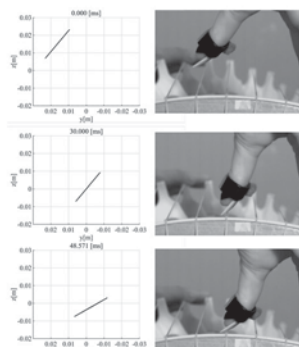


Fig.2: Time response of the vibrational displacement distribution of the pick.

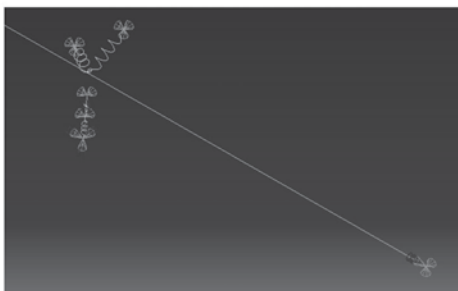
1-1-15

1-1-15 ピアノ弦の空間運動に関する簡易モデル の検討

A simple model for the spatial motion of piano strings

○西口磯春, 落谷颯(神奈川工大), 長沼大介(東京国際大), 若槻尚斗(筑波大)

- ◆報告者らのグループでは、ピアノ音の特徴の1つである2段減衰の原因について検討を行っている。
- ◆この要因として、複数弦の連成とともに挙げられているのが弦の空間運動である。
- ◆この現象の物理モデルとしては、1986年の岩岡と中村らによる理想トランスを含む等価回路によるモデルを先駆けとして、相互誘導トランスモデル、3パラメータモデル等が提案されており、最近の論文でも採用されている。
- ◆前報においては、これらのモデルの関係について示した。本報告では、理想トランスを含む等価回路と等価な質点バネダッシュポットモデルの汎用FEMコードによる計算例を示す。対応する連続体弦モデル(下図)による計算結果も示す。



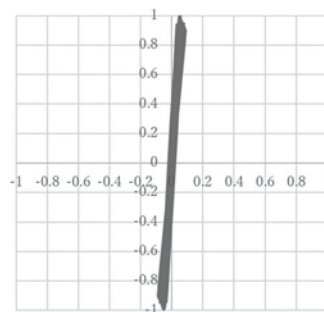
1-1-14

1-1-14 ピアノ弦の空間運動に関する実験的検討

Experimental Investigation of spatial motion of piano strings

☆落谷颯, 西口磯春(神工大), 長沼大介(東京国際大), 若槻尚斗(筑波大)

- ◆2段減衰の要因として、複数弦の連成とともに挙げられているのが弦の空間運動である。
- ◆ハンマー打弦の直後、弦は打弦方向を含む響板と直交する面内で振動するが、時間の経過と共に響板と平行な振動成分が生じ、結果的に弦は空間的に運動する。
- ◆本報告では、回転運動の原因について検討するため、これまでに取得した複数の音域における弦の空間運動の計測データを整理した。
- ◆測定により得られた響板に対する垂直および水平方向の変位について、バンドパスフィルタにより各部分音成分を抽出した。例として、E1弦の基本周波数成分の運動の軌跡を以下に示す。



1-1-16

1-1-16 新型クラヴィコードのソフトウェア音源 に関する検討

Study on software sound source for new clavichord

☆川内詩士, 西口磯春(神工大)

- ◆報告者らは、電気ピアノを対象とする物理モデルの検討を行っている。
- ◆これまでに、ローズピアノの物理モデル音源を JUCE (Jules' Utility Class Extensions) と呼ばれる C++言語用のフレームワークを用いて VST プラグインとしての実装を行った。また、アフタータッチを利用したピッチベンドの実装を行った。
- ◆本研究では、新型クラヴィコード (Fig.1) の物理モデル音源にピッチベンドを実装することを最終目標に、クラヴィコードの物理モデルの文献調査を行うとともに、新型クラヴィコードの物理モデル音源について検討を行った。



Fig1: Neovichord

1-1-17

1-1-17 砂を触り奏でる楽器「SandSynth」の 開発と演奏体験

Creation and performance experience of SandSynth:
a musical instrument that plays by touching sand.

☆伊東天夢, 及川靖広, △菅野由弘(早大理工)

- 本稿では砂を入力装置とした楽器である SandSynth の制作とその音響的評価について論ずる。
- センシングカメラ Intel Realsense とコンピュータを使用
- ド, レ, ミのような離散的な音階ではなく連続的な周波数を演奏可能
- 周波数特性, スペクトログラムを用いた楽器評価

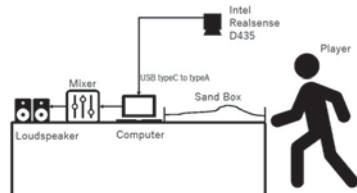


Fig.1 SandSynth の構成

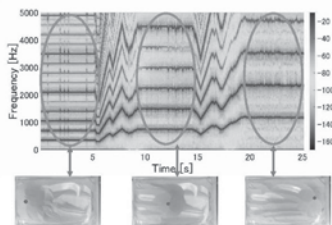


Fig.2 演奏時の周波数変化

1-3-2

1-3-2 この音何デシベル? 小学生を対象としたワークショップ

A What is this sound in dB? -Workshops for primary school children

○上田麻理, 西口磯春, 清原良三(神奈川工科大)

筆者らが大学生や社会人の音響学を学ぶ初学者用に数式を使わずに、音の大きさの理解度を測るための教材を開発した「音の大きさチャート・キット」を用いて沖縄と東京都利島村の小学生に対するワークショップを行ったのでその結果を紹介する。

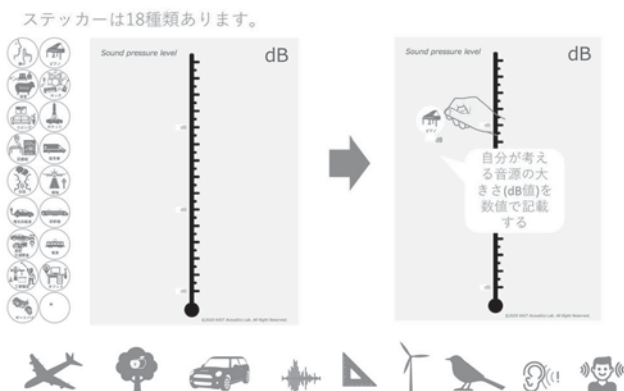


Fig. 「音の大きさチャート・キット」を用いた

小学生へのワークショップ

1-3-1

1-3-1 空間デザイナー向け音響教育のための飲食店の 音環境評価アプリケーションの開発その2

Development of a restaurant sound environment assessment application to provide acoustic education for space designers Part2.

○上田麻理, 田中哲雄(神奈川工科大), 津本祐一(乃村工藝社)

空間デザイナーが音に配慮した空間設計を行うには音響学の知識を学ぶ前に、知識の吸収には自分の感覚など自分事化して、それを客観的に分かりやすく示すことが重要であると考えた。そこで空間デザイナーらとともに音環境の意識付けアプリケーションの開発と評価を行った。続報として今回は残響・響きの感覚に対する訓練アプリケーションを開発した。



Fig. (株)乃村工藝社の空間デザイナーさん達

1-3-3

1-3-3 言語聴覚士の音響学で マスクングをどう教えるか

How to teach "sound masking" in acoustics for speech therapists.

○竹内京子(順天堂大)

- ◆言語聴覚士は、ことばの障害のリハビリを行う職業。言語聴覚士養成校では、音響学(聴覚心理学を含む)が必修科目。聴覚心理学の学習項目に「マスクング」がある。
- ◆しかしながら、音響学・聴覚心理学の授業と聴覚検査実習や補聴器・人工内耳の授業は連携がない場合がほとんどである。それゆえ、現役言語聴覚士に対する調査によると、これらの科目との関係が卒業後も分からない者も多い。科目間の連携が求められる。
- ◆本発表では、「マスクング」の現在の授業内容とその問題点を示し、音響学の授業から、これらの科目、その後の臨床へと興味を持ってもらうためには、今、何ができるかを考える。
- ◆聴覚検査はマスクングをよく知らないと誤った検査値が出る。単なる検査手順だけでなく、マスクングとは何か、どのようなしくみになっているのかについて詳しく知る必要がある。
- ◆マスクングは、一般的に知られている知識ではない。さらに、聴覚検査は、医療関係者以外が経験することはなかなか難しい。しかし、これらを知らずに、マスクングだけを扱うのは難しい。
- ◆今後は、音響学・聴覚心理学担当教師が聴覚検査を知るための機会を作ることを検討し、科目間連携のある、よりよい授業ができる環境を作りたい。

1-3-4

1-3-4 サイン音の提示方向による印象の違いに関する検討

Differences in impressions depending on the direction of sound presentation

◎鮎川菜(九州大・芸工), △湯浅美里, 西平宗貴(三菱電機・デザイン研), 山内勝也(九州大・芸工)

- ◆ 既往研究により, 音を前から提示する場合と後ろから提示する場合は, その印象及び反応に差が見られることが示唆されてきたが, その効果の有無は音源の種類による。
- ◆ 本研究では, サイン音を対象として提示方向による印象の違いについて検討した。また提示方向を正中面だけでなく斜め方向にも増やし, 検討を行った。
- ◆ 前後2方向の評価結果を比較すると, サイン音を刺激とした場合, ほとんどの刺激において印象の差はないことがわかった。
- ◆ 斜め方向の評価を含めた前方対後方の比較においては, 前方からの提示のほうが後方からの提示よりも Valence 値(誘発性)が低く, Arousal 値(覚醒度)が高くなることがわかった。これは, 既往研究の結果と異なり, その原因について今後更なる検討が必要である。

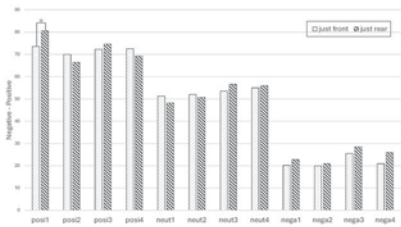


Fig.1: Valence ratings for head-on and head-back presentations.

1-3-6

1-3-6 自動車室内音楽鑑賞時に付与するサイン音の発音方向の制御が検知性と妨害感に与える影響

The Effect of Directional Control of Auditory Signs on their Detectability and Disturbance in Acoustic Environment with Music in Car

○田上宣昭, △磯崎賢太, △石黒留衣, △川田歩(バイオニア), △林真琴(東北バイオニア), 山内 勝也(九州大・芸工)

- ◆ 自動車室内音楽鑑賞時に付与されるサイン音の発音方向の操作が, サイン音の検知性, 妨害感に与える影響について, 主観評価実験を通して検討した。
- ◆ 提示方向条件として Fig.1に示す9つの方向を設定した。(仰角0度: H1~5, 仰角45度: M1~3, 仰角90度: V0)
- ◆ 水平面, 冠状面では横方向から発音されるサイン音は聞き取りやすくなること, 正中面は方向の制御による差はあまりみられないことが観察された。
- ◆ 今回の結果においては, 検知性と妨害感に正の相関がみられ, 発音方向の操作のみでは妨害感を低減させるのは難しいことが分かった。

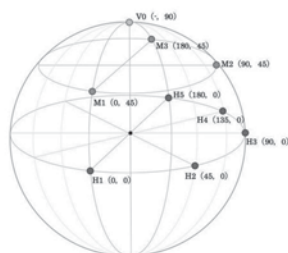


Fig.1: Direction list (horizontal angle, elevation angle)

1-3-5

1-3-5 繰り返しによるサイン音の緊急感の変化に関する検討

Effect of repetition on urgency impression of auditory signs

☆高橋瑞季(九州大・芸工),

浅川香, △陳薇雅, △粟野智治(三菱電機・情報総研), 山内勝也(九州大・芸工)

- ◆ 先行研究によって, 音の繰り返し回数が増すほど緊急感の評価も高くなることが示されているが, その対象は簡潔な音であった。
- ◆ そこで, 周波数構成やリズム変化が複雑な音15種類を対象とし, 繰り返し回数の変化が緊急感に与える影響について検討した。繰り返し周期 500 ms で 2 回もしくは 4 回繰り返した音刺激を作成し, 緊急感と不快感の印象評価を 5 段階で求めた。
- ◆ 各刺激の平均評定値をプロットし, 回数条件の異なる同じ刺激を線分で結んだものを Fig.1 に示す。クラスター分析の結果, 回数の増加で緊急感と不快感が上昇または減少する音, ほとんど変化しない音に分類された。
- ◆ また, ドライビングシミュレータを用いて, 運転状況下での音の印象についても検討した。

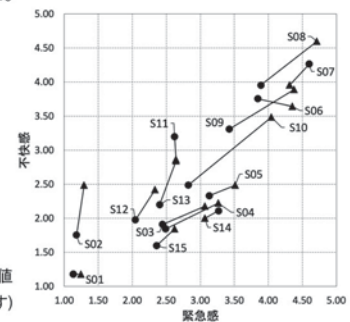


Fig.1 緊急感と不快感の平均評定値 (●は2回, ▲は4回繰り返しを示す)

1-3-7

1-3-7 筆記音が筆記動作へ与える影響: フィードバック操作による検討

The effect of writing sound on writing motion: Investigation by auditory feedback manipulation

☆坂口晃啓, 松井淑恵(豊橋技科大院)

- ◆ 紙とボールペンによる筆記において, 操作した筆記音をフィードバックしたとき筆記動作に影響するかどうかを調査した。
- ◆ ペンタブレットにより筆圧と筆記時間が記録できる環境を構築し, 参加者10名に課題(15種類の文字や図形をランダムに並び替えたもの, 図1)を書き写させた。筆記音はペンタブレットに張り付けたピックアップマイクで収集し, イヤホン経由で聴取させた。
- ◆ 筆記音に操作を加えない non-alter 条件を基本に, 100 ms から 600 ms を 100 ms 刻みの 5 水準で遅延を加える delay 条件, Highpass, Lowpass の 2 水準のフィルタを適用する filter 条件を設定した。
- ◆ 平均筆圧が 100 ms 遅延で低下した。筆記時間は 100 ms, 300 ms の遅延で短く, 400 ms, 600 ms の遅延で長くなる傾向があった。
- ◆ 筆記音の周波数特性の変更は筆記動作には影響しなかった。



Fig.1: Example of an experimental task.

1-3-8

1-3-8 仮想空間内の映像オブジェクトまでの距離に適合する音量

Plausible loudness based on the distance from a listener to a visual object in a virtual space

©岩崎泰士, 長谷川知美, 杉本岳大, 浜本邦基(NHK)

- ◆仮想空間内の映像オブジェクトまでの距離が異なる映像に適合する音量の範囲を調べる主観評価実験を行った。
- ◆評定者は、リスナから評価対象のオブジェクトまでの仮想空間内の距離が2m, 4m, 8mの映像に適合すると感じられる、評価対象のオブジェクトから発せられる音の音量の上限および下限を、フェーダを調節することにより回答した。
- ◆実験の結果 (Fig.1) からコンテンツによって映像に適合すると感じる音量の絶対値は異なるが、上限と下限の平均値の差は9 dB 程度であることが分かった。

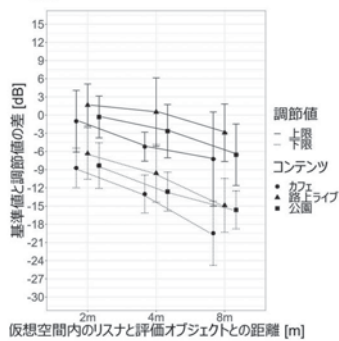


Fig.1: 仮想空間内の映像オブジェクトまでの距離に対する「映像と適合する」と感じる音量の上限と下限の平均値と標準偏差

1-3-10

1-3-10 打法および聴取方法の違いを考慮したバドミントン打音に対する好ましさの印象構造

Impression structure of preference for badminton hitting sounds considering differences in hitting and listening method

○宮入徹, 服部遊(都産技研),

村越弘章, 小澤佳佑, 田中謙介, 千葉 慎一郎(ヨネックス)

- ◆背景: バドミントン競技において、打音は用具購入の動機の一つである。本稿では、打法(ドライブ, スマッシュ)および聴取方法(ヘッドホン聴取, 実打)が、打音の好ましさ印象に与える影響を検討した。
- ◆結果: Table 1の重回帰分析結果(目的変数: 打音の好ましさに対する評価得点, 説明変数: 音質を表す因子の因子得点)より以下の知見を得た。
 - > 打音の好ましさに対して、打法および聴取方法によらず「美的・金属性」が大きく寄与する。
 - > スマッシュ打法はドライブ打法よりも、好ましさに対する「迫力」の影響が増大する。

Table 1 Multiple regression analysis results

(a) Headphones listening and drive-shot		(b) Headphones listening and smash-shot	
変数	標準偏回帰係数	変数	標準偏回帰係数
美的・金属性	0.597	美的・金属性	0.643
迫力	0.352	迫力	0.408
空間	0.276	空間	0.297
調整済 R ² 値	0.80	調整済 R ² 値	0.64

(c) Actual game listening and drive-shot		(d) Actual game listening and smash-shot	
変数	標準偏回帰係数	変数	標準偏回帰係数
美的・金属性	0.698	美的・金属性	0.520
迫力	0.332	迫力	0.462
空間	0.223	空間	0.203
調整済 R ² 値	0.58	調整済 R ² 値	0.60

1-3-9

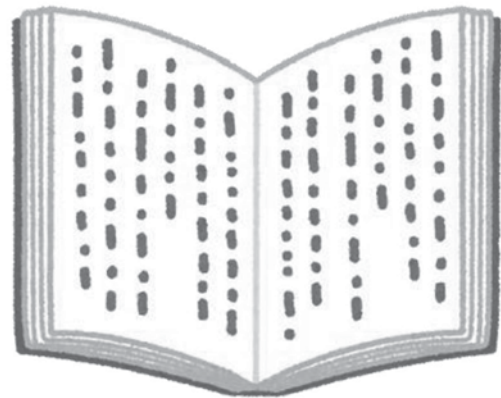
紙をめくる音の音質評価

Sound quality evaluation of the sound of paper turning

☆黒田隼海, 倉本恭次(神奈川工科大),

三浦貴大(産総研), 上田麻理(神奈川工科大)

電子書籍の普及が進んでいる。他方紙媒体の本では、触ったときの感触や音などが本ごとに異なり、それが魅力の一端を担うと考えられる一方、その工学的な制御方法は未確立である。そこで本研究では、紙をめくる時の音に着目し、材質ごとの音響測定・感性評価を行った。



1-3-11

1D-CAE と聴感アンケートの連携によるアジャイルサウンドデザイン

Agile sound design by linking 1D-CAE and simple aural evaluation experiments

○石田康二(Sound One), 柏崎 紘, 村木垂未香, △松本拓也(小野測器), △土居 格(MO Tech Ideas), △後藤泰宏(Sound One)

- ◆思想: 良い音は設計の上流でこそ創られるべき!
- ◆提案: モデルで音生成 ⇒ 聴感評価 ⇒ 物理量との紐づけ ⇒ モデルでパラメータスタディ ⇒ 音生成 のサイクルをアジャイルに循環
- ◆事例: ロードノイズの固体伝搬疑似音を評価。重回帰式より、20~50 Hz 帯域と比較し、63~100 Hz 帯域の POA 値の方が、不快・心地よい 軸の評価値に与える影響が大きい。これをパラメータスタディの一要素とすることで、聴感印象を考慮した設計が可能になる。

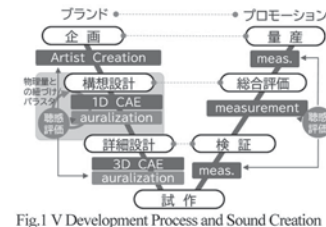


Fig.1 V Development Process and Sound Creation

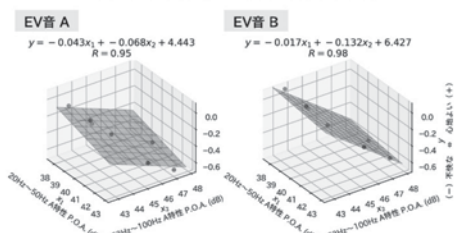


Fig.2 Relationship between evaluation and physical values Multiple regression analysis results

1-3-12

1-3-12 背景騒音および他者の存在が製品音に対する印象に及ぼす影響

The influence of the Presence of ambient noise and another person on the perception of product sound

◎中村史香, 宮入徹, 仏服部遊(都産技研)

- ◆製品の付加価値の一つとして、静音性が様々な製品に取り入れられている。しかし、製品音の印象は使用環境により変化するため、静音性に関する明確な指標が存在しないというのが現状である。
- ◆本稿では、使用環境として背景騒音と他者の存在に着目し、これらが製品音に対する印象に及ぼす影響の把握を目的とする。
- ◆製品音と背景騒音を組み合わせた音を提示した際、他者の存在有無で印象が有意に変化することが確認できた。
- ◆静音性に関する指標として、印象と物理量・心理量の関係も確認したところ、マスクトラウドネスとの高い相関があった。

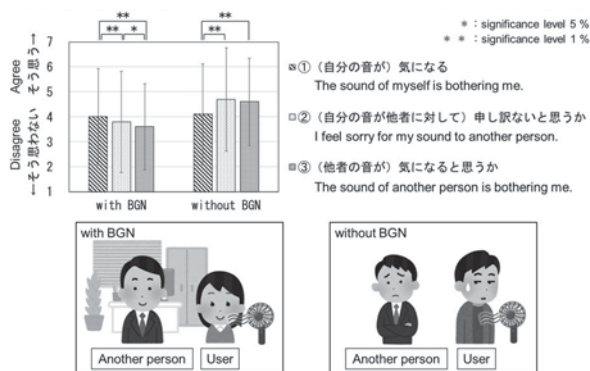


Fig. 1: Averaged evaluation results

1-4-2

1-4-2 イタリア語の促音・非促音に対する発声速度と単語内位置の影響

The effects of speaking rate and syllable position on Italian singleton and geminate

◎天野成昭(愛知淑徳大), 山川仁子(尚綱大), 林良子(神戸大)

イタリア語の促音・非促音の対数時間長による分離・予測の有効性を調べることを目的として、第2、第4音節に/k/の促音・非促音を含む非単語のミニマルペアをイタリア語母語話者5名に様々な速度で発声させ、その音声を解析した。その結果、閉鎖区間の対数時間長と音節の平均対数時間長を用いると、第2音節(Fig. 1a)と第4音節(Fig. 1b)の両方において、促音と非促音のデータが判別境界とほぼ平行に分布し、かつ判別境界と回帰直線の傾きもほぼ同じとなった。さらに判別境界による誤判別率は低く、回帰直線の決定係数は高かった。よって、対数時間長を用いれば、発声速度や単語内位置によらずイタリア語の促音・非促音の分離・予測が可能であるといえる。この結果は、対数時間長が言語非依存の不変的音響特徴の要素であることを示唆している。

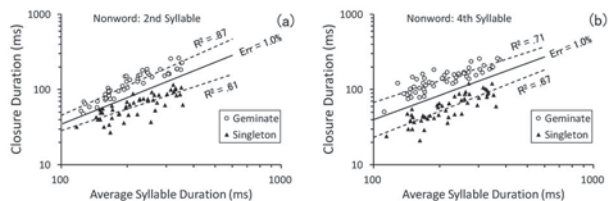


Fig. 1. Scattergram of singleton and geminate consonants at (a) the second syllable and (b) the fourth syllable in nonwords. The solid line represents the discriminant of singleton and geminate consonants. The dashed lines represent the regression equations of singleton and geminate consonants.

1-4-1

1-4-1 対数時間を用いたイタリア語の促音・非促音の不変的特徴

Invariant characteristic of Italian singleton and geminate using logarithmic duration.

◎山川仁子(尚綱大), 天野成昭(愛知淑徳大), 林良子(神戸大)

発声速度の影響を受けない不変的な音響的特徴を特定するため、促音・非促音による意味の対立があるイタリア語を対象に、促音と非促音の予測および判別における対数時間長の有効性について検討を行った。イタリア語母語話者5名が発声した6段階の発声速度による促音・非促音を含む単語を解析対象とし、促音・非促音の無音区間の時間長と平均音節時間長を対数変換して、促音と非促音の判別分析および回帰分析を行った。その結果、促音と非促音の回帰直線と判別境界がほぼ平行となり、誤判別率は1.3%と極めて低かった(Fig. 1)。つまり、対数時間長を用いることで、発声速度に影響されることなく促音・非促音のカテゴリをよりよく表現できるといえる。

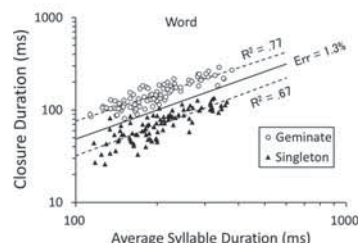


Fig. 1. Scattergram of geminate and singleton consonants. The solid line represents the discriminant of geminate and singleton consonants. The dashed lines represent the regression equations of geminate and singleton consonants.

1-4-3

1-4-3 日本語を母語とする英語学習者による強勢の音声的実現: ピッチパターンの特徴

Phonetic realization of English stress by Japanese learners: analysis of pitch pattern

◎北原真冬(上智大), 田嶋圭一(法政大), 米山聖子(大東文化大)

日本語を母語とする英語学習者(JE)をTOEICスコアにより2群に分け、高レベル(JEH)8名と中レベル(JEM)8名、および統制群として英語母語話者(AE)8名に、強勢位置が移動する派生語の組(例: computer - computation)、計83単語を発音させた。強勢段階(第1強勢, 第2強勢, 無強勢)が異なる隣接音節のピッチの差を計測し、話者グループ(AE, JEH, JEM)による違いを検討した。その結果、JEMは第1強勢の後のピッチを大きく下げるという点で、日本語ピッチアクセントの高低のパターンを英語の強勢にあてはめていることが示唆された。それに対し、第2強勢から無強勢音節へのピッチ変化はどの話者グループも小さく、その差も有意ではなかった。また、JEM, JEHともに無強勢音節から第1強勢にかけてAEよりもピッチを大きく上昇させることも判明した。

LME analysis of pitch difference between neighboring syllables by language groups

	1-0 transition		2-0 transition		0-1 transition	
	Est.	t-value	Est.	t-value	Est.	t-value
Intercept	-19.060	-3.067**	-2.361	-1.403	-14.965	-2.516*
JEH	-4.209	-0.500	-0.284	-0.129	20.832	2.546*
JEM	-21.847	-2.589*	-0.902	0.407	29.221	3.570**

*p < .05 **p < .01

1-4-4

1-4-4 ドイツ語接尾辞にみられる鼻腔破裂

—日本人習熟度別学習者の生成分析—

Nasal plosion on German suffixes in the unstressed syllables by German natives and Japanese German learners

○粕谷麻里乃(東邦音大・音), 荒井隆行(上智大・理工)

- ◆背景: ドイツ語接尾辞「子音+ <en>」は弱化の影響を受けやすい。弱化により、母音は縮約し、子音部分は伸長し、成節化へと移行する。同時に「鼻腔破裂 (nasal plosion)」を伴う音声変異の可能性がある。
- ◆目的: ドイツ語音声現象として鼻腔破裂を調査する。母語話者と日本人上級学習者を対象に子音別 (有声・無声破裂音) に出現を調査した。
- ◆結果: 母語話者は、英語と同じ /l/, /d/ に加え、/p/, /k/ にも鼻腔破裂を伴う発話を確認した。閉鎖開放時に spike が観察され、この時に口蓋帆の下降し、鼻腔へと開放させた。有声破裂音 /b/, /g/ は確認できなかったが、話速により子音弱化を確認した。上級学習者は /l/, /d/ のみであった。
- ◆考察: ドイツ語では、予想以上に日常発話で鼻腔破裂を発生させて弱音を処理していると考えられる。英語と同じ解釈では説明できない。

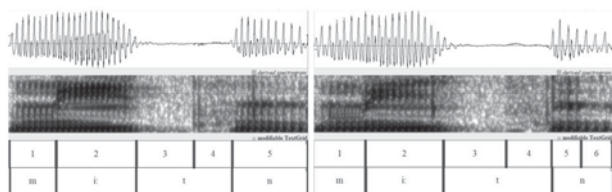


Fig. 1: Examples of phonetic annotation using Praat. (Left: with nasal plosion, Right: without nasal plosion)

1-4-6

1-4-6 中国語における「そり舌音」の調音方法について

Articulation of "retroflex consonants" in Mandarin Chinese

◎呉琪(筑波大), 孫静(神戸大), 北村達也(甲南大)

- ◆中国語には特有な子音 zh (/tʂ/), ch (/tʂʰ/), sh (/ʃ/) があり、日本人学習者にとっては特に困難である。
- ◆従来の研究や多くの教科書はこれらの子音を「そり舌音」と定義しているが、実際の発音時において、舌先をそらせているかどうかに関して、実験的調査は限られている。
- ◆本稿では、MRI を用い、中国語母語話者の発話における母音に先行する「そり舌音」の調音方法を考察した。
- ◆中国語の母音に先行する「そり舌音」の調音時には舌先をそらさず、「舌の前方を持ち上げる」という調音法を用いたことが分かった。
- ◆本稿の実験協力者による「そり舌音」の調音位置は後続母音の影響をほとんど受けていないことが見られた。

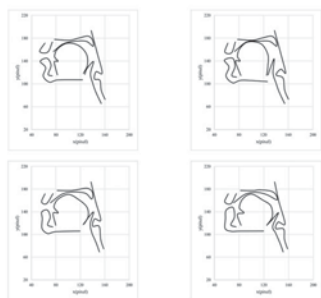


Fig. 1: Contours of /tʂ/ (top left), /tʂʰ/ (top right), /ʃ/ (bottom left), /ʒ/ (bottom right) by a native Chinese speaker.

1-4-5

1-4-5 中国語母語話者による日本語・中国語読み上げ文の非流暢性

Disfluency in Japanese and Chinese Read Speech by Chinese Native Speakers

◎孫静, 林良子(神戸大)

- ◆本研究では、中国語母語話者による日本語と中国語の読み上げ文における非流暢性の特徴を分析した。日本語の「北風と太陽」と中国語の対応するテキストを用い、19歳から22歳の中国語母語話者20名(男性6名、女性14名)に読み上げてもらった。
- ◆日本語の読み上げ文では、欠落、付加、阻止(途切れ)、繰り返し(音、モーラ、音節、語の部分)、延伸、フィラーといった非流暢性が観察された。特に、初級レベルの日本語学習者である実験協力者は、日本語の音連続を正確に発音できないことが多く、途切れや繰り返しが頻繁に見られた。
- ◆中国語の読み上げ文では、複雑な単語や不慣れた語彙に対する代用や付加の言い間違いが観察され、長文や複雑な文において途切れが多く見られた。
- ◆両言語の非流暢性の平均生起数を比較すると、外国語である日本語の方が母語である中国語よりも非流暢性が多く出現し、特に途切れが顕著であった。一方、フィラーの生起数は少なく、文読み上げ時にフィラーをほとんど使用しないという共通点が見られた。

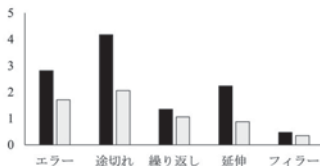


Fig. 1: Average frequency of Disfluency in reading Chinese and Japanese tasks

1-4-7

1-4-7 F0 推定に寄与する声道の輪郭成分の検討

Investigation of contour components of the vocal tract contributing to F0 estimation

◎脇田真子, 竹本浩典(千葉工大), 平井啓之, 前川喜久雄(国語研)

- ◆ソース・フィルタ理論では、基本周波数 (F0) は声帯の緊張や肺圧で制御され、声道形状の制御とは独立とされる。
- ◆われわれは前報で正中矢状断面における声道の輪郭から音響特徴量を予測して音声合成を行い、音韻性や抑揚を再現可能であることを示した。これは、ソース・フィルタ理論に反する結果と言える。
- ◆本研究では、声道の輪郭から F0 を予測するモデルを再構築し、説明可能な AI を用いて、F0 推定に寄与する声道の輪郭成分を検討した。
- ◆Integrated Gradients を用いて F0 上昇・下降に寄与する声道の輪郭成分の可視化を行なった結果、各母音に共通して F0 上昇時は喉頭が前上方に、下降時は後下方へ移動した。また、喉頭蓋の輪郭が F0 推定に大きく影響することが明らかになった。舌骨を前方に牽引する動作や、喉頭を下降させる動作は声帯張力を変化させて F0 に影響するため、これらの動作が喉頭蓋部分に現れていると考えられる。

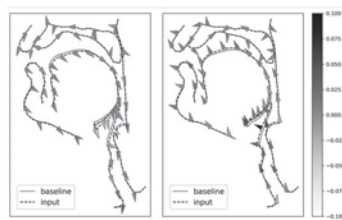


Fig. 1: Visualization results of contour components contributing to F0 estimation. (left: /a/ 120 Hz, right: /a/ 200 Hz)

1-4-8

1-4-8 rtMRI で捉えた F0 制御に関わる 調音動作の分析

Analysis of Articulatory Movements Related to F0 Control
Captured with rtMRI

◎大谷祐人, 澤田隼, 大村英史, 桂田浩一(東京理科大)

我々は real-time MRI (rtMRI)で収録した調音運動データを入力とする音声合成法を提案してきた。この手法ではrtMRI 動画のみを入力とし、F₀や有声/無声を含めて高い精度で音声を復元しているが、rtMRIは声帯振動を捉えることができないため、ソース・フィルタモデルに直感的に反する結果といえる。real-time MRI で収録した調音運動データからは、F₀変化と連動する微小な運動が観察された。本研究では男女各2名に話者を増やし、rtMRI 動画からの音声合成を通じて同様にF₀を得られるか確認する。また、オプティカルフロー推定を用いて運動を可視化することで、F₀制御及び有声/無声の制御に関わる調音動作を分析する。

実験の結果、rtMRI 動画から合成した音声の F₀ RMSE は話者別に 24.5 Hz から 26.3 Hz であり、いずれの話者についても十分に高い精度で F₀ が得られることを確認した。オプティカルフローによる分析では、すべての話者について共通して、F₀ 上昇に伴い喉頭から舌を前方・上方に引き上げる動作があり、F₀ が特に高い領域では舌骨付近を中心として前転するような動作が見られた。男性話者については F₀ が低い領域で別の F₀ 制御が行われることを確認した。また、男性話者に限り一貫して無声子音の発声時に喉頭から舌を上方に引き上げる動作が見られた。

今後は一部話者の特に低い F₀ 領域で見られた動作について、個人差や声帯に及ぼす影響を検討する予定である。また、有声/無声の違いがどのように現れるかについて、さらなる調査を行う予定である。

1-4-10

1-4-10 方言音声の獲得・学習

Japanese dialects of spoken language and acquisition/learning

○白勢彩子(東京学芸大)

- ◆日本語は方言のバリエーションが豊かで、音声にも各地域の特色がある。日本語の方言研究の歴史は長く、多くの成果が蓄積されてきている。しかしながら、発達過程においてどのように方言が形成されるかについてはあまり触れられてきていない。本稿では、方言研究の流れを概説し、方言の獲得・学習に関する研究をとりあげて、今後の研究の発展を促したい。
- ◆方言研究は大きく分けて、共時的な変異(バリエーション)とその要因を明らかにしようとする立場と、共時的な現象を扱いつつも通時的な変化を論じようとする立場との二様がある。
- ◆現代日本語では共通語化の問題が大きく、多くの地域で方言は「危機言語」である。しかしながら、一律に共通語化したように見えても、方言が完全に失われたわけではなく、方言と共通語が使い分けられている、つまり方言と共通語とが共存している状況にある。さらに、単に両者が切り分けて使われるのではなく、中間段階の言語が形成されていることも確認されている。
- ◆このような新しい方言や共通語と方言との「バイリンガル(バイダイアレクタル)」な状況が各人においてどのような過程で形成されたのかについては非常に重要な観点と考えられるが、発達過程における方言の形成に着眼した研究は、ごく限られている。
- ◆大規模なデータの集積が難しい領域であり、こうした研究がどのような意義を持つのかについても考えたい。

1-4-9

音声信号から調音状態への逆問題における 非一意性に関する検討

Investigation of non-uniqueness in the inverse problem
from speech sounds to articulatory states

☆梶浦一真, 脇田真子, 竹本浩典(千葉工大), 平井啓之,
前川喜久雄(国語研)

- ◆音声から調音状態への逆問題では非一意性が存在すると考えられているが、英語における先行研究では多くの場合一意に定まると報告されている。
- ◆本研究では、リアルタイム MRI を用いて ATR503 文を朗読中の調音運動を撮像してフレームごとに舌の輪郭を抽出し、同時に録音した音声から音響特徴量を抽出した。そして母音ごとに音響特徴量に対して Kmeans 法でクラスタリングを行ったのち、同一クラスタに属する舌輪郭形状に対して MeanShift 法でクラスタリングを行った。
- ◆Figs. 1, 2 で示すように、音響的には同一のクラスタに属する舌輪郭形状が、形状的には複数のクラスタ(黒色の破線と灰色の実線)に分かれたことから、複数の舌形状から同一の音が生成されている非一意性が見られた。

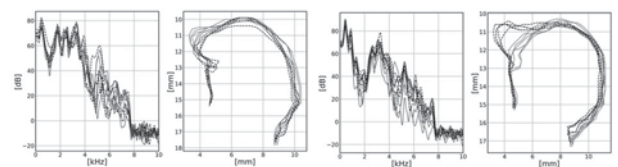


Fig. 1: power spectra (left) and tongue shapes (right) for /e/ Fig. 2: power spectra (left) and tongue shapes (right) for /o/

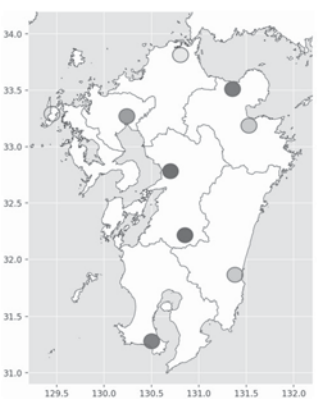
1-4-11

1-4-11 九州諸方言における母音融合の 生産性の検証を目指して

Towards productive varification of vowel coalescence in Kyushu Japanese

○松浦年男

- ◆九州地方の音声・音韻は地域による多様性があるが、ほん研究ではその中で母音融合の生産性と地域差に注目して研究を行った。
- ◆先行研究では、名詞の種類という語彙的な条件による生起頻度の違いが指摘されている。
- ◆コーパス調査の結果、代名詞と与格の組み合わせでの母音融合の生起率に地域差が確認された(右図)。
- ◆辞典・方言記述書の分析では、母音による偏りや名詞の種類による影響が見られ、特定の母音や名詞で融合が多いことが示された。
- ◆フィールドワークのデータからも、特定の名詞や母音による母音融合の違いが観察され、生起条件に地域差があることが示唆された。
- ◆今後の課題として、名詞+与格の意味役割の違いが母音融合に与える影響を詳細に検討することや、用法基盤モデルを適用して母音融合のメカニズムを明らかにすることが必要であると指摘した。



1-4-12

1-4-12

音声の方言らしさのエッセンス

Essence of the dialectal character of speech.
○内田照久(大学入試センター 研究開発部)

目的 イントネーションの高低を反転した逆相 f_0 パターンの音声から「未知の方言印象」が想起される。なぜ想起するのかを、基本的な正弦波状 f_0 パターンの音声を手がかりにして検証。

実験 1. 標準語の正相 f_0 音声から逆相 f_0 音声を生成して評価。
2. 正弦波状 f_0 パターンの正相・逆相、及び、その振幅値と周波数を操作した音声も生成して評価。

「標準語らしさ」、「地方の方言らしさ」などを10件法で評定

結果 首都圏の国立大学1年生256名による評価の結果、原抑揚の逆相 f_0 や正弦波状 f_0 の音声は、言語音声としては破綻せず、地方の方言として認知された。しかし、 f_0 の変化が速かったり、遅かったりすると、方言印象の想起は消失した。この結果から、1.5 Hz 近傍の変動成分に方言らしさの情報が含まれる可能性。

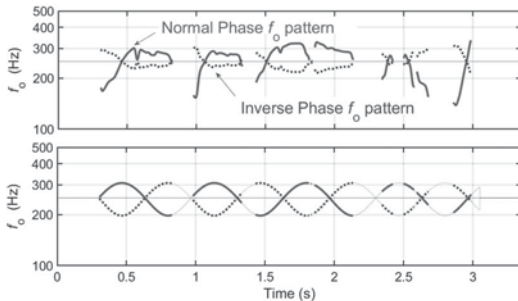


Fig. 1 Examples of original intonation and sinusoidal f_0 patterns.

1-4-14

1-4-14

男性による「両声類」の音声の研究

A Research on the "Ryousei-Rui voice" of Male voice

○高野佐代子、長塚全
(金沢工大, Zen Voice Factory)
Sayoko TAKANO, and Zen Nagatsuka

- ◆ 近年、男性による女声のような発声（あるいはその逆）をエンターテインメントの一つとして行う人々が増加し、主に動画投稿サイトなどを通じて「両声類」の名で広く知られるようになっている。
- ◆ 声道長と基本周波数を独立に操作した合成音声を用いたSD法による聴取実験の結果、3因子が得られ、それぞれ素適さ、女らしさ、強さと命名した。
- ◆ 地声から声道長だけ短くする(スペクトルを高くする)と女らしさが若干上がるが、素適さが顕著に下がる。基本周波数の値を高く、かつ声道長を短くすると、女らしさと素適さが上がる。

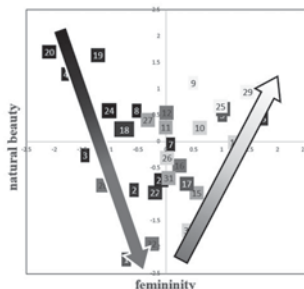


Figure 1. The results of the hearing experiment

1-4-13

1-4-13 会話場面による日常会話音声の韻律の多様性

Diversity of prosody of everyday speech for conversation situation
○石本祐一(ものづくり大/国語研)

- ◆ 多様な種類の日常会話を収録した大規模なコーパスである『日本語日常会話コーパス』(CEJC)を用いて、様々な会話場面における韻律の違いを調べた。
- ◆ 韻律の要素のうち声の高さと発話速度に着目し、音響特徴量として発話の平均F0と発話の平均モーラ長を用いる。
- ◆ その結果、声の高さに関しては、子どもや配偶者といった家族に対しては低い声で、友人には高い声で話していることがわかった(Fig. 1)。また、会話場面に同席している参加者の属性によっても声の高さが変わることが観察された。
- ◆ 発話速度に関しては、発話の向け先による違いはわずかであった。

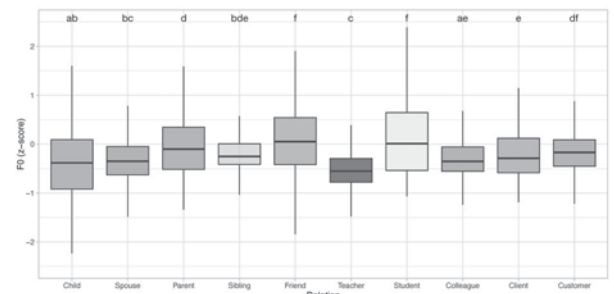


Fig.1: Distributions of the average fundamental frequency of utterance for relationships between speaker and addressee.

1-4-15

1-4-15 応答表現「それな」の使用実態とその音調の分析

-日本人大学生による生成調査の結果報告-

Prosodic features of "sorena" in response expressions by Japanese university students

☆朴 彩恩, 鮮于 嫻(埼玉大)

- ◆ 本研究は、応答表現で**若者言葉の一つである「それな」**を対象とし、その使用実態とその音響的特徴の把握を試みた。本稿は、日本人大学生を対象とした生成調査の結果を報告する。
- ◆ 分析対象は154の音声データで、「上昇調」、「非上昇調」の二つのグループに分類し、F0値を求めた。その結果、平均値として、123件の「**上昇調**」は**267.6 Hz**、31件の「**非上昇調**」は**234.6 Hz**であった。
- ◆ 音調幅については、受け入れの程度の大きさの関与が示唆された。「上昇調」の場合、音調幅は「52.7 (同意しない) → 100.4 (どちらでもいい) → 357.9 (とても同意する)」となった。「非上昇調」の場合、「4.1 (同意しない) → 6.1 (どちらでもいい) → 16.3 (とても同意する)」のように**同意の程度による差が見られた。**

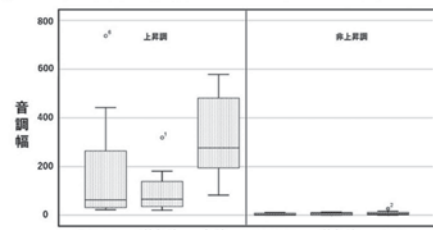


Fig.1: Comparison of tone width by degree of agreement between "rising tone" and "descending tone"

1-4-16

1-4-16 EMA データを中間情報とする IPA 系列からの音声合成

Speech Synthesis from IPA Sequences Using EMA Data as Intermediate Information.

☆丸山虹樹, 澤田隼, 大村英史, 桂田浩一(東京理科大)

本研究では、調音情報を豊富に含んでいる IPA から、調音運動データの EMA データを経由して、音声を生産する手法を提案する。提案モデルを Fig. 1 に示す。提案モデルは4つのモジュールから構成されている。EMA generator は、IPA を入力として EMA を生成する。Integration block は、生成された EMA データを IPA データと結合する。Mel generator は、IPA と EMA データを入力として、メルスペクトログラムを生成する。hifi-GAN は、メルスペクトログラムから音声を生産する。

IPA と EMA データ、音声を含んだ日本語のデータセットを用いて評価したところ、EMA データを中間情報として使用することで、音声の品質が向上することを確認できた。特に、EMA データに含まれている調音位置が調音点となる音に関しては、EMA データを用いる有効性を確認することができた。

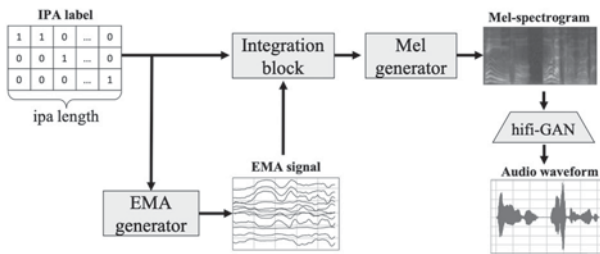


Fig. 1 Structure of speech synthesis model from IPA through EMA.

1-4-18

1-4-18 超音波・MRI 調音データの輪郭自動抽出の試み—DeepLabCut を用いて

Attempts to automatic extraction of contours of articulatory organs from ultrasound and MRI data using DeepLabCut

◎孫静(神戸大), 北村達也(甲南大), 林良子(神戸大)

◆言語学習や発音矯正の研究では、舌の動きの解析が重要であり、特に超音波と MRI による調音データは詳細な舌の動きを捉えるために広く利用されている。しかし、これらのデータから舌の輪郭を自動抽出するのは技術的な課題が多く、手動解析は時間と労力がかかる。

◆従来の輪郭抽出手法 (Edgetrak, GetContours, DeepEdge など) は手動描画が必要であり、精度に制約がある。近年、機械学習技術の進展により動画解析の自動化が進んでいるが、超音波と rtMRI の両方を対象とした研究は限られている。

◆本研究では、動物の行動解析のために開発された DeepLabCut を用いて、超音波と rtMRI のデータから舌の輪郭を自動抽出する方法を試みた。大量かつ多様な超音波および rtMRI データセットを用い、モデルをトレーニングすることで、不明瞭な輪郭やノイズの多いデータでも高精度な抽出が実現された。

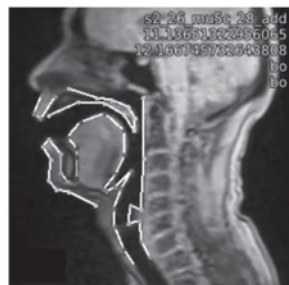


Fig.1: Contours of speech organs extracted from rtMRI image

1-4-17

1-4-17 複数の声道立体形状の分析に基づく歌手手のフォルマントの生成メカニズムの検討

Examination of mechanism for generating the singer's formant by acoustic analysis of three-dimensional vocal tract shapes

☆志々目樹, 戸田菜月, 竹本浩典(千葉工大), 高橋純(大阪工大)

◆前報でわれわれは、「歌手手のフォルマント (Singer's Formant: SF) は主声道 (喉頭腔を除く声道) に由来する第3フォルマント (F3)、喉頭腔に由来する第4フォルマント (F4)、梨状窩に由来するピークが集約したものである」という仮説を立てた。

◆本研究ではテノール歌手1名の5母音と他の3名の男性歌手の/a/の歌唱発声時の声道立体形状 (声道モデル) を MRI で計測して分析することで、仮説を検証した。

◆音声スペクトル (SPC)、声道モデルの伝達関数 (VT)、主声道モデルの伝達関数 (VTp) を分析した結果、仮説は必ずしも常に成立せず、SFの周波数、SFを構成するピークの種類や数は歌手によっても母音によっても異なっていた。

◆SFを構成しているピークの由来は、F3は主声道の共鳴、F4は喉頭腔の共鳴であるが、これらと近接する可能性がある5番目のピークの由来は、主声道、梨状窩、咽頭の横モードの3通りがあった。

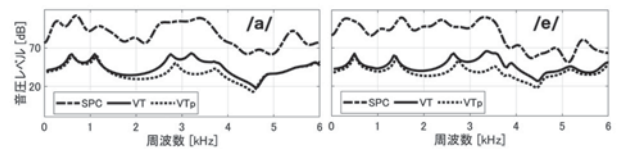


Fig. 1 テノール歌手1名の母音/a/と/e/のSPC, VT, VTp

1-4-19

1-4-19 rtMRI 動画への歯列補填の精度を向上させる手法の検討

Investigation of methods to improve the accuracy of tooth superimposition on rtMRI video frames

☆常盤朔也, 竹本浩典(千葉工大), 前川善久雄(国語研)

◆前報でわれわれは調音運動を記録した rtMRI 動画へ歯列を補填する手法を開発し、多話者でその精度を分析した。

◆補填する歯列周辺と補填先の歯列周辺に矩形領域を設定して位置合わせする必要があるが、その設定が精度に大きな影響を与える。そこで、本研究では、以下の設定を検討した。

- 上顎の提案法1: 口唇を含まないように矩形領域を拡大
- 上顎の提案法2: 前鼻棘周辺のみ限定した矩形領域を設定
- 下顎の提案法: 下顎骨の海綿骨の周辺のみ限定した矩形領域を設定

◆上顎でも下顎でも提案法は従来法に比べて誤差の平均値と最大値、参考値超過率は低下し、精度が向上した。

Table1 従来法と提案法の上顎の補填精度

	話者	s1 s4 s20 s21			
		平均値	最大値	参考値	参考値超過率
従来法	平均値	6.0	3.8	4.6	2.6
	最大値	35.1	27.3	28.2	25.0
	参考値	15.3	8.7	15.7	11.3
	参考値超過率	2.6	3.9	0.2	1.4
提案法1	平均値	5.2	2.2	4.3	1.2
	最大値	30.1	16.6	16.4	12.9
	参考値	15.3	8.7	15.7	11.3
	参考値超過率	1.6	0.4	0.0	0.0
提案法2	平均値	5.6	3.8	3.5	1.5
	最大値	28.9	10.4	15.3	11.7
	参考値	15.3	8.7	15.7	11.3
	参考値超過率	1.2	0.1	0.0	0.0

Table2 従来法と提案法の下顎の補填精度

	話者	s1 s4 s20 s21			
		平均値	最大値	参考値	参考値超過率
従来法	平均値	13.3	6.8	12.5	4.0
	最大値	33.9	21.1	37.1	25.1
	参考値	17.5	11.5	12.8	13.0
	参考値超過率	14.4	2.9	47.4	0.8
提案法	平均値	8.0	5.5	2.4	1.2
	最大値	29.8	11.4	13.5	12.6
	参考値	17.5	11.5	12.8	13.0
	参考値超過率	0.8	0.0	0.0	0.0

1-4-20

1-4-20 Wav2lip モデルのファインチューニングによる rtMRI 動画の生成

Generation of rtMRI movies by fine tuning of Wav2lip model

☆陸鳴宇(甲南大院), 北村達也(甲南大), 前川喜久雄(国語研)

◆本研究では、音声と顔画像(動画)を学習して作成されたリップシンク動画生成用の深層学習モデルを活用し、このモデルをファインチューニングすることによって音声から rtMRI 動画を生成する技術を提案する。ファインチューニングには『リアルタイム MRI 日本語調音運動データベース』の構築のために収集された rtMRI データを利用する。本研究により生成 AI の応用範囲がひろがることを期待している。

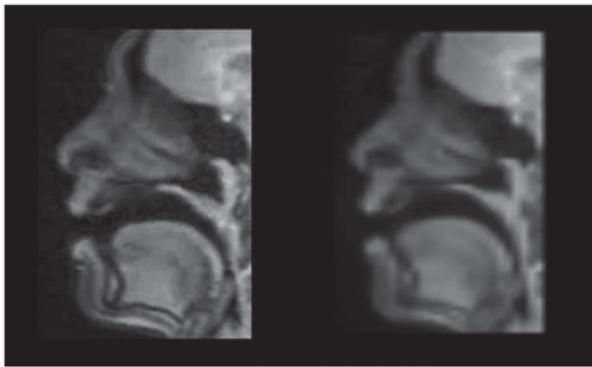


Fig.1: Example of the for the original (left) and generated (right) images.

1-5-2

1-5-2 An Evaluation of 3D Printing for Creating Dummy Heads to Study Crosstalk Cancellation in Bone Conduction

© IRWANSYAH, Sho OTSUKA, and Seiji NAKAGAWA (Chiba Univ.)

- ◆ This study evaluates the use of 3D printing for creating dummy heads to study crosstalk cancellation in bone conduction. Dummy heads were 3D-printed using ABS material for the skull and silicone skins ranging from very soft (OO-32) to slightly hard (OO-74), as shown in the right panel of Fig. 1.
- ◆ The experimental setup employed a pair of B81 bone transducers and PCB 352a24 accelerometers, with the Filtered-X Least Mean Square (FxLMS) algorithm applied to achieve crosstalk cancellation at the mastoid. The cancellation was confirmed at the mastoid and successfully extended to the cochlea, though with decreasing effectiveness at the cochlea location.
- ◆ This crosstalk reduction at the cochlea correlated well with the hearing threshold improvements observed in human subjects, particularly with harder silicone skins, highlighting the potential of these models for bone conduction research.

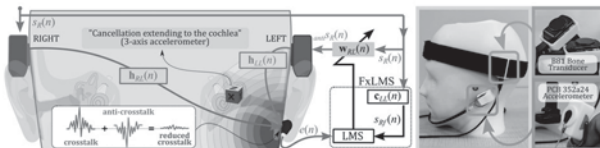


Fig. 1 Illustration of the crosstalk cancellation setup using 3D-printed dummy heads with B81 bone transducers and PCB 352a24 accelerometers.

1-5-1

1-5-1 聴覚フィルタバンクを用いた音質評価指標の計算モデルの応用

Application of computational model of sound quality metrics with auditory filterbank

◎磯山拓都, 木谷俊介, 鶴木祐史(北陸先端大)

- ◆ 【目的】我々のグループが構築した音質評価指標の計算モデルを紹介するとともに、これらの計算モデルが、一般的な応用先として使われる環境音の分析に適用できるか調査する。
- ◆ 【構築した音質評価指標の計算モデル】
Fig.1に音質評価指標の計算モデルのGUIを示す。
 - ラウドネスモデル: Moore-Glasbergのラウドネスモデルに基づいて、時間領域で計算するモデルとして構築
 - シャープネスモデル: Auresのラウドネス依存のシャープネスモデルをもとにして構築
 - ラフネスモデル: Daniel-Weberのラフネスモデルをもとにして構築
 - 変動強度モデル: ラフネスモデルをもとにして、変動強度に合わせたパラメータを用いて構築
- ◆ 【環境音の分析】
 - 構築したモデルの推定結果と従来手法の推定結果は同様の傾向を示した。このことから、我々の計算モデルは、環境音の評価に適用できた。

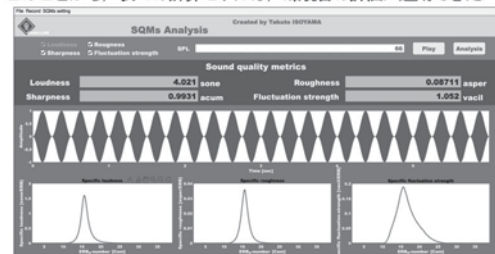


Fig.1: Computational model of sound quality metrics. Download URL: <https://github.com/Isoyama94/Computational-Models-of-SQM.git>

1-5-3

1-5-3 Comparison between bone conduction and behind-the-ear cartilage conduction – the influence of placements

OXiuyuan Qin, Sho Otsuka, and Seiji Nakagawa (Chiba Univ.)

- ◆ A novel type of headphone which transmits sound signal by cartilage conduction (CC) was proposed. The new position which is behind the ear (BTE). This study investigates the feasibility of utilizing (BTE-CC) by comparing their performance to bone conduction (BC). Measurements were taken at three positions: the mastoid process, the squamous part of the temporal bone, and behind the ear. The study measured hearing thresholds, ear-canal sound pressures (ECSP), and acceleration using a piezoelectric vibrator at these positions.
- ◆ BC demonstrated better hearing thresholds than CC. The trade-off between higher ECSP and lower vibration in CC leads to higher hearing thresholds, indicating that BC is more efficient in transmitting sound to the inner ear, enhancing hearing perception.



Fig. 1. The measured vibrator and three positions where the vibrator was put.

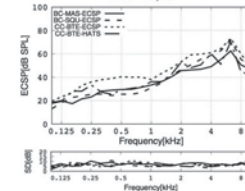


Fig. 3. The results of ECSP measurements and HATS simulation of the vibrator at three different positions

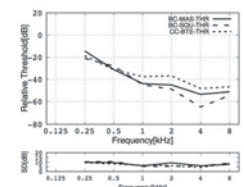


Fig. 2. The results of hearing threshold measurements of the vibrator at three different positions.

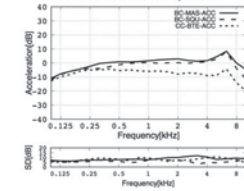


Fig. 4. The results of acceleration measurements of the vibrator at three different positions.

1-5-4

1-5-4 耳介のサイズと硬度が軟骨伝導検出閾に及ぼす影響:成人と小児の比較

Effects of auricular size and hardness on detection threshold of cartilage conduction: Comparison between adults and children

☆田村 茜, 大塚 翔(千葉大), 小谷博子(東京未来大), 中川誠司(千葉大)

- ◆骨伝導の装用性の問題を解決する方法として“軟骨伝導”が提案され、補聴器やスマートホン、イヤホンなどに応用されている。
- ◆軟骨伝導は、外耳道への挿入を嫌う乳幼児のための“貼る”イヤホンとしても応用されているが、乳幼児の耳介は大人と比較して小さく柔らかく、聞こえに差異が存在する可能性がある。
- ◆5~9歳の小児を対象として軟骨伝導の検出閾、耳介のサイズおよび硬度の計測を行い、それらのパラメータの相関の調査および成人の結果との比較を行った。
- ◆Fig. 1 に検出閾を示す。成人に比して小児の検出閾が低下した。250, 2000 Hz では耳介サイズ検出閾間に正相関が認められた。一方、小児の検出閾が上昇した。500, 1000 Hz では耳介硬度が高いほど検出閾が低下した。
- ◆周波数によってその特性に違いが生じるものの、耳介のサイズと硬度が軟骨伝導知覚に影響を及ぼし、小児と成人の間に知覚の差異を生じさせることが示唆された。

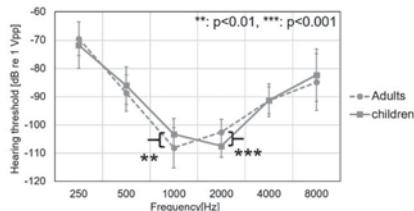


Fig.1: Hearing thresholds for the cartilage conduction.

1-5-6

1-5-6 軟骨伝導補聴器購入時の公的な支援について

Public support to purchase cartilage conduction hearing aids

○西村忠己, 細井裕司, △北原 紘(奈良医大)

- ◆軟骨伝導補聴器の購入時に得られる公的な支援の現状を把握するため、取扱施設となっている38都道府県122施設にグーグルフォームを使用したアンケート調査を行った。
- ◆36都道府県,77施設(63.1%)から回答が得られた。
- ◆身体障害者の有無、年齢、外耳の状態、両耳か片耳かで分類し、各条件での結果を都道府県ごとに集計し、日本地図上に示した。
- ◆公的な支援が得られるかどうかは、地域差を認め、その偏りは、各条件とも同じような傾向を認めた。
- ◆等しく公的な支援が得られるように、行政への働きかけと軟骨伝導補聴器の規格化が重要であると考えられた。

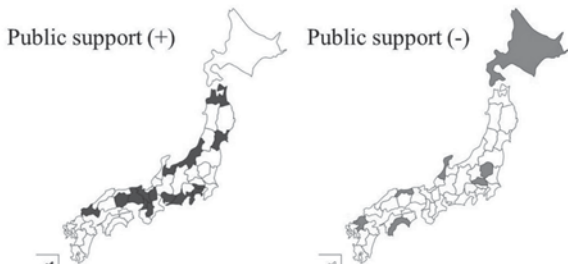


Fig.1: Regional localization of public support record for cartilage conduction hearing aids in the children with hearing disability certificate.

1-5-5

1-5-5 多様な分野での軟骨伝導の普及

Promotion of cartilage conduction in various fields

○細井裕司(奈良医大)、下倉良太(大阪大)、西村忠己(奈良医大)

- ◆軟骨伝導は細井が2004年に発見した気導とも骨伝導とも異なる第3の音伝達経路である。
- ◆国際科学誌にCartilage conductionの論文は40編以上となり学問的には認知された。
- ◆新しい音伝達経路によって、今まで存在しなかったジャンルが切り開かれ、製品化がはじまった。ステレオサウンドで「水中音楽」を楽しむ新音楽ジャンル、宝石で音を聞くサウンドジュエリー、高齢者が窓口でしっかりと聞き取れる窓口用軟骨伝導イヤホン、多数の人が同時に軟骨伝導音を聞ける軟骨伝導集団音響システム、軟骨伝導イヤホンによる耳ツボの刺激などである。
- ◆Fig.1に水中音楽鑑賞のイメージ図、Fig.2にサウンドジュエリーを示す。



Fig.1: Underwater music

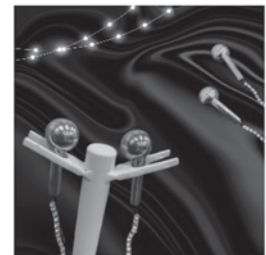


Fig.2: Sound jewelry

1-5-7

1-5-7 軟骨伝導集団音響システムの開発

Development of cartilage conduction mass sound system

○犬伏 雅士(奈良女子大, 大阪大), 細井 裕司(奈良医大)

- ◆人の聴覚経路については次の3つが知られている。500年以上前から知られていた気導や骨伝導に対し、軟骨伝導は2004年に細井が発見した第3の聴覚経路である。
- ◆軟骨伝導音を個々の使用だけでなく、集団で一度に聴取するシステムを開発した。Fig.1に軟骨伝導集団音響システムを示す。
- ◆学校現場で、聴覚学習の新たな指導方法や実験装置として、軟骨伝導集団音響システムを理科教育に利用できる事を示唆している。小学校と高校の実践をFig.2やFig.3に示す。

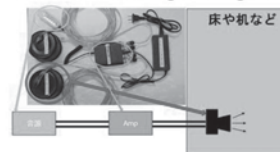


Fig.1 Cartilage conduction mass sound system



Fig.2 Practice with pupils and parents in an elementary school

- ◆軟骨伝導デバイスが、2025年に行われる大阪・関西万博に採用される。



Fig.3 Practice in a high school

1-5-8

1-5-8 世界初の軟骨伝導による両耳聴インカムの開発

Development of the world's first bilateral hearing intercom using cartilage conduction

○中川雅永、△河田晋輔、△矢崎賢吾 (CCHサウンド)

- ◆世界で初めの両耳で聞ける軟骨伝導インカムを開発した。
- ◆現在ほぼすべてのインカム (intercom) は、片耳装用であり、片耳で通信音を聞き、もう一方の片耳で周囲の音を聞く。
- ◆その欠点は通信音を片耳で聞くため、両耳加算効果が得られず確実な聴取が困難なことがある。また、周囲の音を片耳で聞くため周囲の音の音源定位ができず音源の方向や距離がわからない。
- ◆軟骨伝導のイヤホンを用いれば、両耳で通信音を聞くことができ、同時に周囲の音も把握できる。
- ◆有料老人ホームの職員に試聴していただき、従来インカムと比較した結果、通信音や周囲音の聞き取りやすさで軟骨伝導インカムが勝っていた。
- ◆Fig.1に軟骨伝導イヤホン装用を示す。



Fig.1: Cartilage conduction earphone for intercom
 Communication sound: Heard in both ears (clearly audible)
 Ambient sound: Heard in both ears (full perception of sound direction and distance)

1-5-10

1-5-10 帯域雑音除去を達成する軟骨伝導二次音源を用いた Least Mean Square

Least Mean Square using secondary cartilage sources to achieve band noise cancellation

○加藤良樹、下倉良太(大阪大)、笹岡直人(鳥取大)、飯国洋二(大阪大)

- ◆オープンイヤーでのノイズキャンセリングを目指しており、オープンイヤーで音の制御が可能な軟骨伝導に着目した。
- ◆本研究は、音声と狭帯域や広帯域のノイズが重畳した環境で、二次音源として軟骨伝導から信号を出力することで、鼓膜での狭帯域ノイズの除去を目標としており、本発表では、シミュレーション実験に関して発表する。
- ◆消音を可能とするFiltered-x least means square (FxLMS) アルゴリズムに、ノイズの周波数を逐次推定する適応バンドパスフィルタを組み込んだ ABF-FxLMS アルゴリズムを提案した。
- ◆消音効果を高めるため、更新の変化量を定めるステップサイズを可変ににする Variable Step Size (VSS)アルゴリズムを組み込んだ。
- ◆提案アルゴリズムを用いた結果、ノイズを狭帯域ノイズでは20dB、広帯域ノイズでは19dB 消音されることが確認できた。

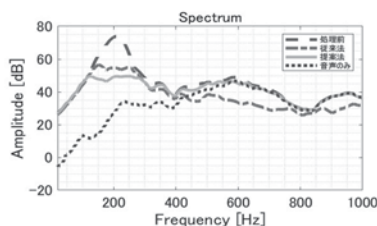


Fig.1: Spectra of four signals (narrowband noise at 200 Hz)

1-5-9

1-5-9 Delayed-X harmonics synthesizer による狭帯域ノイズ制御を目的とした軟骨伝導アクティブノイズコントロール

Delayed-X harmonic synthesizer to minimize narrow bandpass noise by cartilage conduction secondary source

○下倉良太、東美幸、飯国洋二(大阪大)、笹岡直人(鳥取大)

- ◆軟骨伝導音を二次音源とし、外部から到来する騒音を相殺するオープンイヤー型アクティブノイズコントロールを目指す。
- ◆同時に外部から到来する音声を無加工で鼓膜に届けるため(選択的消音)、Delayed-X harmonic synthesizer という正弦波出力のアルゴリズムを利用して、複数の正弦波で狭帯域ノイズのみの消音を試みた。
- ◆ノイズの帯域に5つの正弦波を配置し、各正弦波のステップサイズパラメータを狭帯域ノイズのスペクトル包絡の大きさに合わせて調整することにより、スペクトルピークを削るような滑らかな消音(-11.3dB)を達成した (SERSS-DXHS)。スペクトル包絡の推定には10次の線形予測符号 (LPC) を用いた。

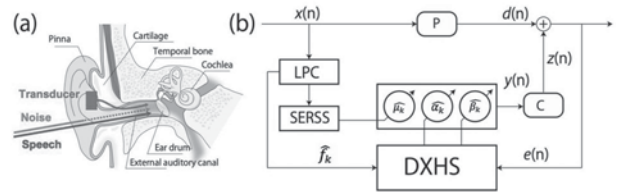


図 (a) 軟骨伝導を用いたオープンイヤー型アクティブノイズコントロール、(b) 対象信号 $x(n)$ (騒音+音声) から LPC で騒音のスペクトル包絡を推定、包絡応答のステップサイズで DXHS を適応 (SERSS-DXHS) ; P は一次経路、C は軟骨伝導の経路

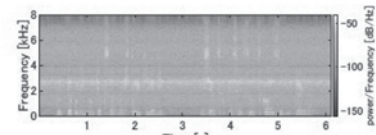
1-5-11

1-5-11 適応ラインエンハンサ及び軟骨伝導による選択的消音システム

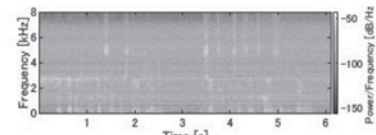
Selective noise control using adaptive line enhancer and aural cartilage

○笹岡直人、△吉永康佑(鳥取大)

- ◆騒音環境下の会話において外耳道を塞がずに騒音を消音し、音声のみを聴取する選択的消音システムを提案する。
- ◆提案システムは、外耳道内における環境騒音の消音に能動騒音制御を用いており、1次経路である気道に対して、2次経路として軟骨伝導を利用する。
- ◆能動騒音制御システムの入力信号として、消音対象である広帯域騒音と音声成分の分離が必要であるため、一つのマイクロホンで分離可能な適応ラインエンハンサを用いる。
- ◆Fig.1に示す計算機シミュレーション結果より、外耳道によって強調された周波数帯の騒音が消音されていることが確認された。



(a) Proposed system outage



(b) Proposed system running
Fig.1: Spectrogram of error signal.

1-5-12

1-5-12 ブラインド音声処理で拓く骨伝導音の未来

The future opening for body/bone-conducted speech using blind speech processing

○中山仁史 (広島市大院)

- ◆骨伝導は気導雑音の影響を受けにくいので、98dB SPL(-20dB SNR)の環境下でも骨伝導として発声を得ることができる。
- ◆“話す骨伝導”は系統的な信号計測や評価が難しく、これまでの知見や各研究成果をベースに比較することが難しい研究である。
- ◆“話す骨伝導”は骨伝導音のみで音質改善できるかが課題である。
- ◆ブラインド音源分離より音素や音声記号を単位としたブラインド音声処理を着想した。
- ◆ブラインドによる音声処理が実現できれば、骨伝導や軟骨伝導へ拡張することで減衰した周波数特性の改善や応用が期待できる。
- ◆高磁場高騒音環境下の骨伝導光音のみを用いた erSS-NMF で MRI の雑音抑圧と高周波数成分の改善を確認した。
- ◆身体内を伝搬する様々な臓器が発する骨伝導音を対象としたブラインド医療/医用音源分離なども近い将来に拓くことが期待できる。

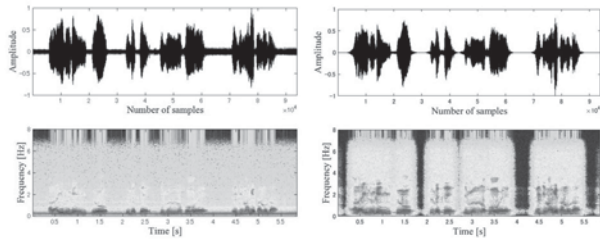


Fig. 1: Optical BCS with MRI noise

Fig. 2: Retrieval sound of Optical BCS

1-6-2

1-6-2 音声特徴を制御可能な潜在変数をもつ HSMM 構造化アテンションに基づく音声合成

Controllable speech synthesis with structured attention based on hidden semi-Markov models

◎藤本崇人, 橋本佳, 南角吉彦, 徳田恵一(名工大)

本稿では、HSMM 構造化アテンションに基づく音声合成において、潜在変数系列を基本周波数やエネルギーなどの外部ツールで抽出された補助特徴量と紐づけることにより、潜在変数の解釈性と制御性を向上させる。また、補助特徴量の精度をモデル化することで、抽出エラーを考慮し、外部ツールに対する依存性を軽減する。客観評価実験やサンプル例から提案モデルの制御性の向上を示した。

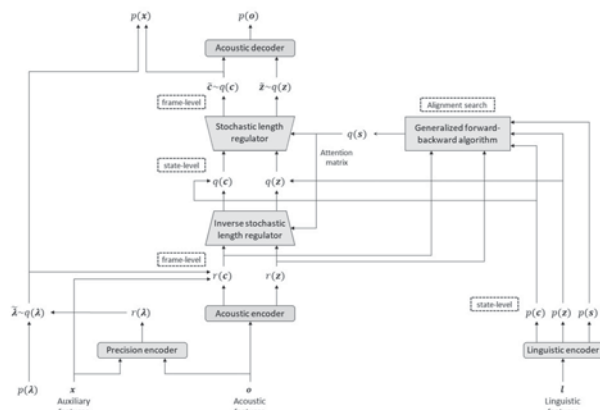


Fig. 1: Structure of the proposed model.

1-6-1

1-6-1 Physics-informed Neural Networks に基づく音声合成

Speech Synthesis based on Physics-informed Neural Networks

◎横田和哉, 小倉正崇, 坂爪孝熙, 藤本一毅, 阿部雅二郎 (長岡技術科学大学)

- ◆近年の機械学習の進歩により、ニューラルネットワークを用いて数値解析を行う手法が提案されている。その一つである Physics-Informed Neural Networks (PINNs) は、支配方程式による拘束を損失関数に導入することで、物理的に満足する解を得る手法である。
- ◆PINNs は逆問題に対して有用な手法として、音響分野においても研究が進んでいるが、PINNs を用いて音声合成を行った例は報告されていない。
- ◆本稿では、PINNs を用いた音声合成(順解析)と、音声データからの声門流速波形の同定(逆解析)について報告する。
- ◆提案法は逆解析のための複雑なアルゴリズムを構築する必要なく様々な逆問題に適用可能であり、簡便であるという長所がある。

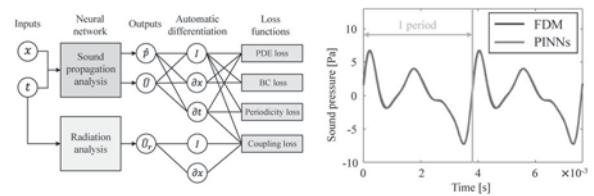


Fig. 1: Structure of PINNs for voiced sound synthesis.

Fig. 2: Synthesized voiced sound (vowel /a/).

1-6-3

1-6-3 音素埋め込みスキップ接続を用いた継続長拡張に頑健な音声合成

Duration expansion robust neural text-to-speech with phoneme embedding skip connection

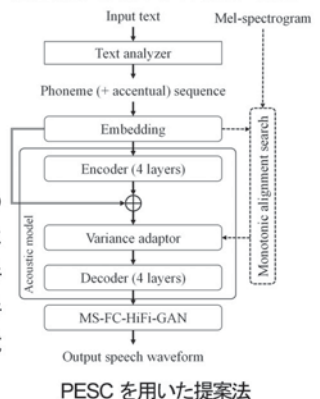
◎小椋志志, 岡本拓磨, 大谷大和 (NICT), 戸田智基 (名大/NICT), 河井恒 (NICT)

- ◆目的: ニューラルネットワークを用いたテキスト音声合成 (TTS) システムにおける継続長制御精度の向上
- ◆課題: 従来の TTS モデルはエンコーダ出力特徴量において過去・未来の音素情報が混在し音声特徴が不明確化
→ 音素継続長拡張時の音声品質低下が課題

◆提案法:

1. 音素埋め込みスキップ接続 (PESC): エンコーダ出力特徴量への音素レベル情報の直接追加による継続長制御精度向上
2. Monotonic Alignment Search 注意機構の対象位置の変更: 音素レベル情報のより正確な保持
3. ガウス継続長モデリング (GDM) との組み合わせ: 各音素の弾性を考慮した継続長制御の実現

◆実験結果: 日本語と英語の TTS システムを用いた実験において、提案手法は短縮条件 (GDM なし) および伸長条件 (GDM あり) において従来方式よりも高品質な生成を実現



1-6-4

1-6-4 話者匿名化したデータを用いる 多話者テキスト音声合成

Multi-speaker Text-to-speech Training with Speaker Anonymized Data.
OHuang Wen-Chin¹, WU Yi-Chiao², 戸田智基¹ (¹名古屋大学, ²Meta)

本研究は、話者匿名化処理を施した音声データを用いた多話者テキスト音声合成モデル学習に関する調査を行う。実験的評価により、本枠組みにおいて有効な合成音声自動評価指標を明らかにする。

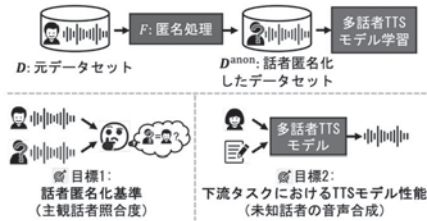


Fig. 1. Problem formulation and goals of this work.

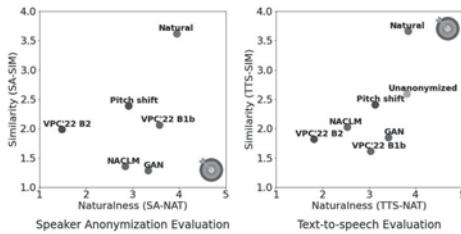


Fig. 2. Scatter plots of the subjective evaluation results. The goal icons are located at the ideal score positions.

1-6-6

1-6-6 感情音声合成のための アラインメント手法の比較

A comparative study of text-speech alignment methods for emotional speech synthesis

©蓮実拓也, 白旗悠真, Welly Naptali, 山本龍一(LINE ヤフー)
Eunwoo Song (NAVER Cloud), 橋健太郎(LINE ヤフー), Jae-Min Kim (NAVER Cloud)

- ◆自己回帰型テキスト音声合成 (text-to-speech: TTS) の学習にはアラインメントを要するが、正確なアノテーション付与はコストがかかる。
- ◆外部の強制アライナーの利用や教師なしアラインメント学習は、コストの問題を回避するが、それぞれ音素セットの整合性や学習の不安定さといった問題が残る。また、感情音声合成のような発展的なTTSタスクでは、各手法の有効性は調査されていない。
- ◆本稿では、感情音声合成における各アラインメント手法の比較を行う。また、教師なしアラインメント学習の安定化のために、少量の正確なアノテーションデータと大量のアノテーションなしデータが得られる状況に適した、半教師ありアラインメント学習を提案する。

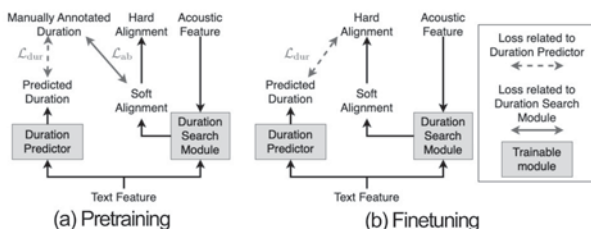


Figure 1. Proposed semi-supervised alignment learning

1-6-5

1-6-5 複雑な感情音声合成のための 潜在空間でのデータ拡張

Emotional Latent Space Augmentation for Speech Synthesis with Mixed Emotions

☆野崎雄斗, △加藤大地, △廣岡聖司(東京大院), △金崎朝子(東京工業大)

- ◆目的: 複数の感情を組み合わせた混合感情を扱う音声合成
- ◆課題: 混合感情音声合成は未だに難しい
 - 混合する感情を音声データにラベリングするコストが高い
- ◆提案法: 感情埋め込み空間でのデータ拡張とそれを用いた学習法
 - 感情の構造モデルに基づく感情埋め込み空間でのデータ拡張
 - ◇ 埋め込み混合, 埋め込み強度増幅, 埋め込み反転とノイズ加算
 - 話者埋め込み空間でのデータ拡張法に触発された、拡張された感情埋め込みを用いた学習
 - ◇ 出力音の感情埋め込みと入力埋め込みに一貫性を持たせる
- ◆実験結果: ESD テストデータで性能を既存手法と比較
 - 混合感情音声合成における音声の自然さの客観評価の改善
 - MOS と感情識別の主観評価で既存手法と同程度の性能
- ◆今後: 埋め込みの一貫性に加え感情の識別結果の一貫性も考慮

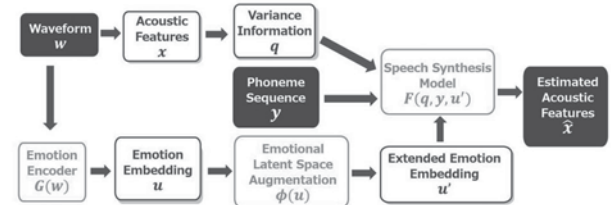


Fig1. Proposed Training Framework

1-6-7

1-6-7 A Study on Replacing Self-Attention with Linearly Recurrent RWKV in Speech Encoders

☆ Wen Shen TEO, Yasuhiro MINAMI (UEC)

- ◆Explores replacing Self-Attention in conventional streaming speech encoders with the linearly recurrent component of RWKV.
- ◆Used Finch – a newer RWKV that supports dynamic recurrence.
- ◆Evaluated against chunkwise Zipformer on LibriSpeech and CSJ, at 0.32s, 0.64s, and 1.28s chunk sizes and 1.28s, 2.56s, and 5.12s left context lengths.

Results

- ◆RWKV achieves comparable performance for small chunk sizes and limited left context, but with increasing context Self-Attention outperforms.

Discussion

- ◆We posit: Differences in text and speech modalities necessitate adaptation of RWKV to outperform conventional Self-Attention in streaming ASR.

1-6-8

1-6-8 事前学習モデルを用いた感情埋め込みによる発話スタイル依存型話者照合

Utterance-Style-Dependent Speaker Verification Using Emotion Embedding by Pre-trained Models

☆高山響, 西田昌史(静岡大), 柘植覚(大同大), 黒岩眞吾(千葉大)

これまで我々は、なりすましに頑健な話者照手法として「感情を想定した発話スタイル依存型話者照合」を提案し、評価してきた。本研究では提案手法の照合性能を向上させるため、事前学習モデルから感情埋め込み、話者埋め込みを抽出し、それぞれから得られる照合スコアを重みづけ和により統合するアプローチを提案する。また、ユーザ評価実験を通じて、感情演技発話によってユーザにかかる負担を基にシステムの利便性について分析した。Table1より、提案手法はBaselineである話者照スコアのみからの従来手法に比べて照合精度を改善することができた。また、Fig.1よりニュートラルの発話に比べて、テスト回数を重ねることで感情を再現することへの負担が軽減されることが明らかになった。

Table 1 EER of each method [%]

	Emotion				Average
	Anger	Joy	Sadness	Neutral	
Baseline	1.28	2.16	1.19	1.76	1.60
ECAPA-TDNN	1.04	1.84	1.07	1.40	1.34
wav2vec 2.0	0.87	2.00	0.86	1.45	1.30
HuBERT	0.88	1.75	0.82	1.05	1.13

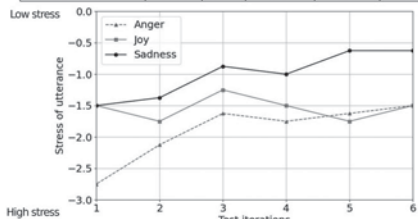


Fig. 1 Stress of emotional utterances when the stress of neutral utterance is set to 0

1-6-10

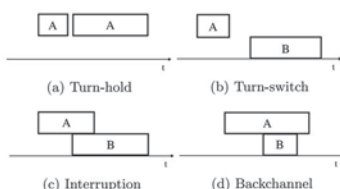
1-6-10 End-to-end ダイアライゼーションのための自然な会話データシミュレーション

Improving the Naturalness of Simulated Conversations for End-to-End Neural Diarization

◎山下夏生, 堀口翔太, 本間健(日立製作所)

- ◆End-to-end ダイアライゼーション (EEND) では、アノテーションされた実データの量は限られているため、大規模なシミュレーションデータで事前学習し、次にターゲットドメインの実データで適応するという手法が一般に用いられている。
- ◆シミュレーションデータは EEND の学習に重要であるにも関わらず、最適なシミュレーション手法については十分に調査されていない。
- ◆そこで、話者交替 (turn-taking) の概念を取り入れ、4 種類の遷移タイプを考慮しながら前から順番に組み合わせるシミュレーション手法を提案する。
- ◆CALLHOME データセットと CSJ データセットを用いた評価では、提案手法がより自然な会話シミュレーションデータを作成し、EEND の精度を向上させることを確認した。

Fig.1: Four utterance transition cases in two-speaker scenario.



1-6-9

1-6-9 複数話者音声からの話者特徴量抽出

Speaker embedding extraction from multi-speaker recordings

◎堀口翔太, 安藤厚志, 森谷崇史, 佐藤宏, 芦原孝典, 俵直弘, デルクロアマーク(NTT)

- ◆話者特徴量の利用シーンは話者識別や照合にとどまらず、近年は複数話者音声に対するアプリケーション (話者ダイアライゼーションや目的話者音声処理) においてよく用いられる。一方で、発話重複の存在により、複数話者音声からの話者特徴量抽出には困難が伴う。
- ◆本稿では、話者特徴量抽出器のプーリング層において用いられるアテンション重みを再帰的に計算することで、複数話者音声に含まれる話者それぞれに対する話者特徴量抽出を実現した。
- ◆話者照合および話者ダイアライゼーションにおける評価により、提案手法の有効性を示した。またアテンション重みの可視化により、提案手法が音源分離に相当する機能を獲得していることが示唆された。

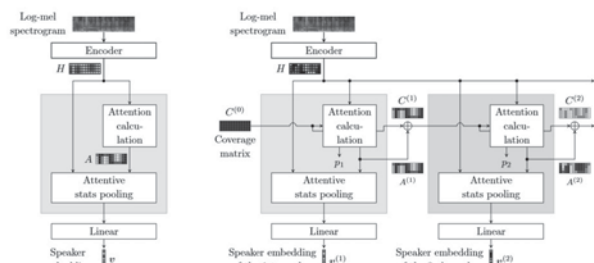


Fig. 1: Basic speaker embedding extractor from single-speaker recordings (left) and the proposed extractor from multi-speaker recordings (right).

1-8-1

1-8-1 定在波音場中の鉄球に作用する音響放射力の球径依存性の評価

Evaluation of diameter dependence of acoustic radiation force acting on iron spheres in standing wave field.

◎丸目勝斗, △成田憲一(愛工大), 畑中信一(宇都宮大), 鎌倉友男(電通大), 豊田昌弘(本多電子), 小塚晃透(愛工大)

- ◆超音波振動子 36 個よりなる集束音源を製作し、2 音源を対向させて形成した定在波音場で、超音波浮揚の研究を行っている。
- ◆定在波音場中の音圧分布は、外形 1mm の注射針を導波管として用いるニードル型マイクロホンを試作して測定した。音軸上の音圧の節と腹が検出されることを確認した。Fig. 1 に音軸上の音圧分布を示す。
- ◆定在波音場に 1 mm~7 mm の鉄球を糸で吊して投入し、その荷重の変化を測定した。音源間の位相を変化したところ、Fig. 2 に示す結果を得た。音圧の節が上方にある場合は軽くなり、下方にある場合は重くなることに示され、超音波による力の変化が測定された。

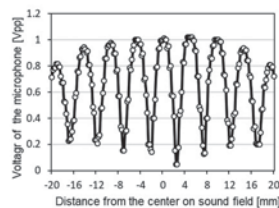


Fig. 1. Sound pressure along the sound beam axis.

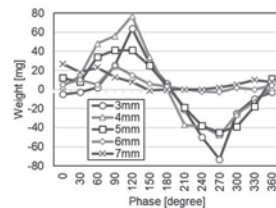


Fig. 2. Weight changes of the balls as a function of the phase difference.

1-8-2

1-8-2 液滴の超音波浮揚における操作範囲の測定

Measurement the operating range in ultrasonic levitation of droplets

☆井上雄大, 畑中信一(宇都宮大), 鎌倉友男(電通大),
△豊田昌弘(本多電子), 小塚晃透(愛工大)

- ◆小型超音波スピーカ 36 個を凹面形状に配置した音源ユニット向かい合わせに2つ配置し, 集束する定在波音場を形成する。
- ◆2つの音源の駆動周波数に0.1 Hzの差を与え, 10秒に1周期間隔で音場を連続的に変化させることで浮揚物体を操作する。
- ◆浮揚物体を上下に動かした際に, それが落下した地点を落下(上), 落下(下)とする。中心から離れた位置で浮揚させた水滴を中心に近づけた際に, それが破裂した地点を破裂(上), 破裂(下)とする。
- ◆Fig. 1に, 浮揚させた水滴の操作範囲を示す。印加電圧が27.5 Vpp以上の時, 電圧を上げることで上下方向に操作範囲が広がるが, 水滴が破裂し操作できない範囲の方がより広がるため, 実質的な水滴の操作範囲は狭くなることがわかる。

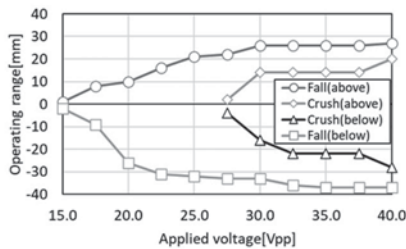


Fig. 1 Operating range of water droplets.

1-8-4

1-8-4 平面上に置かれた穴付き円筒に作用する超音波浮揚力の検討

Investigation of ultrasonic levitation force acting on a cylinder with a hole placed on a plane, by FUJIOKA, Yuta¹, WANG, Yimeng¹, TAMURA, Hideki² and AOYAGI, Manabu¹ (1Murooran Institute of Technology; 2Tohoku Institute of Technology).

☆藤岡夕大¹, Wang Yimeng¹, 田村英樹², 青柳樹¹(¹室蘭工大・院, ²東北工大)

- ◆円筒が微かに浮揚する現象について, 円筒底に座ぐりを設け, その直径を変化させて浮揚力の解析と浮揚実験を行った。
- ◆Fig. 1は座ぐり直径を変化させたときの浮揚を開始した振動振幅値の測定結果である。浮揚開始時の振動振幅値が変化していることから円筒下の空隙が浮揚力に影響することが確認できた。
- ◆Fig. 2に円筒を5 μm 浮揚させた状態で座ぐり直径を変化させたときの円筒に働く力の解析結果を示す。浮揚力は常に正の値となるため, 座ぐり直径によらず浮揚する。土台-円筒間に隙間がある場合は解析と実験で同様の結果とが得られ, 座ぐり直径が浮揚力に影響することが分かった。

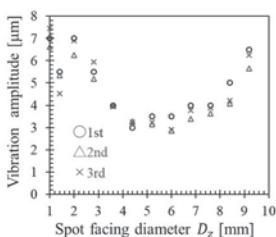


Fig. 1 Minimum vibration amplitude vs. spot facing diameter D_2 required for levitation.

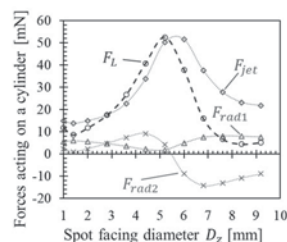


Fig. 2 Forces vs. facing diameter D_2 .

1-8-3

1-8-3 超音波ゲルレンズの透過光波面解析

Analysis of the transmitted light wavefront of the ultrasound gel lens

☆三木陽斗, 中村光佑, 原田裕生, 小山大介(同志社大)

- ◆超音波ゲルレンズは, 音響放射力によってレンズの表面形状を変化させることでその焦点位置を制御できる。
- ◆機械的可動部を持たないため小型・薄型化が可能。
- ◆レンズの透過光波面を解析し, 焦点距離と球面収差の評価を行った。
- ◆印加電圧の増加に伴い, ゲルの表面形状は空気側に向かって凸状に変形した (Fig. 1)。それに伴い, 焦点距離は負値から正値へ変化し, 球面収差は電圧振幅と共に増加した (Fig. 2)。

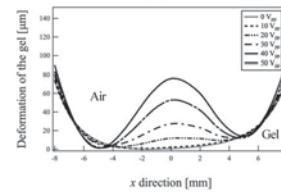


Fig. 1 Surface profiles of the ultrasound gel lens under ultrasound excitation.

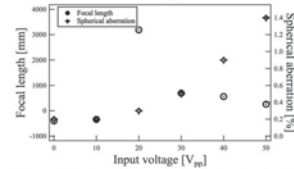


Fig. 2 Focal length & Spherical aberration vs input current.

1-8-5

1-8-5 Generation of acoustic streaming by a cylinder with a cavity located away from vibrating surface.

© Yimeng Wang, Manabu Aoyagi (Murooran Inst. Tech.)

- ◆Using PIV experiments and FEA simulations, the effect of positioning the cylinder with a cavity away from the vibrating surface and closer to the target object was investigated to enhance acoustic streaming at the resonance frequency of 28.2 kHz.
- ◆Through PIV experiments, actual acoustic streaming distribution was obtained as shown in Fig. 1. Although streaming velocity at the same location was enhanced by duralumin cylinder, as shown in Fig. 2, the effect was not significant enough.
- ◆Using FEA simulations, the gap and the cylinder surrounded by the absorbing boundary were considered for eigenfrequency analysis. The distribution and magnitude of the sound pressure at eigenfrequency were used as a criterion to optimize the cylinder dimensions. The sound pressure distribution of original and optimized cylinders under the vibration source are shown in Fig. 3.

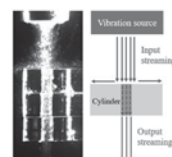


Fig. 1 Photo and schematic of streaming distribution of streaming velocity V_{max} vs. distance from the vibration surface h .

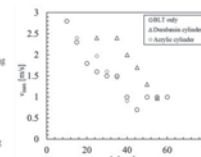


Fig. 2 Maximum output streaming velocity V_{max} vs. distance from the vibration surface h .

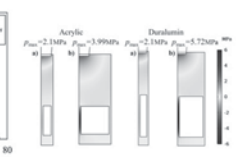


Fig. 3 Sound pressure distributions of a) original cylinders and b) optimized cylinders.

1-8-6

1-8-6 印加超音波音圧レベルに対する浮き上がり火炎特性の検討

Study of floating flame characteristics in relation to applied ultrasonic sound pressure level

○松岡和輝(大同大), △飯田大智, △山本青空, 白石裕之

- ◆これまでの研究結果から、浮き上がり火炎に対して超音波の印加、印加音源の台数の増加、印加する角度間隔の適正化によって燃焼持続時間の増加は確認できていた。
- ◆そこで本研究では、各印加ケースにおいて浮き上がり火炎の燃焼持続時間と音圧レベルの関係性について検討を行う。
- ◆具体的には、浮き上がり火炎に印加する超音波音源を1~3台、角度間隔毎に超音波周波数 21.0 kHz と 23.0 kHz、電圧 1.5 kV ~ 3.0 kV に調整して音圧レベルの測定を行った。

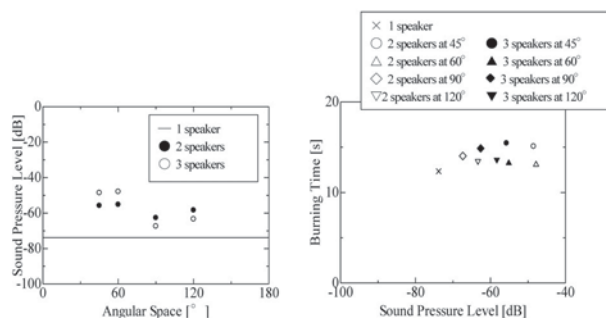


Fig.1: Sound pressure level by various angular intervals at 3.0 kV and 23.0 kHz

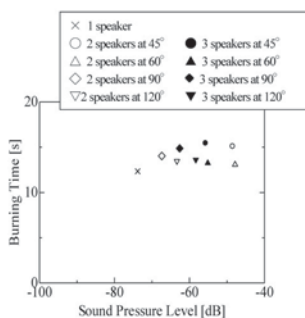


Fig.2: Duration time at 3.0 kV and 23.0 kHz

1-8-8

1-8-8 圧電高分子トランスデューサを用いた送受信兼用型方式による物質内部の透過法について

Through-transmission method inside a material by using one transmitter/receiver piezoelectric polymer transducer

○高橋真幸(山形大学)

これまで、筆者らは P(VDF/TrFE) トランスデューサを用いた、MHz 帯域での空中超音波送受信による物質内部(主として集積回路/IC)の透過画像形成を行ってきた。MHz 帯域を用いる理由は、高分解能の画像を得るためである(高周波数に伴い短波長となり、分解能は向上する)。しかし、高周波数になるに連れて、媒質(空気)や観測対象物により、超音波エネルギーの減衰値は極度に増大する。加えて、超音波送受信システムで発生するノイズの影響も無視できない(受信波がノイズ中に埋没する現象も発生する)。このため、高周波での物質を透過する超音波送受信の場合は、ノイズの影響の少ない、2つのトランスデューサを用いる送受信分離型透過方式が一般的である。本研究では、1つのトランスデューサを用いた送受信兼用型透過方式(Figure 1)を試みたことを報告する。

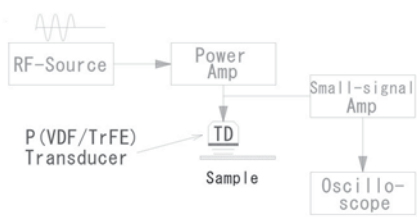


Figure 1 shows ultrasonic system to transmit-receive by the one P(VDF/TrFE) transducer.

1-8-7

1-8-7 屈折と球面収差を考慮した集束超音波計測方法による高感度化

Focused ultrasonic measurement method taking into account refraction and spherical aberration for higher sensitivity

◎田中雄介, △柘植延啓, △野地正明, △小倉幸夫(ジャパンプローブ)

- ◆集束超音波探触子により試験体内部欠陥を探傷する場合、通常は図1のように送信探触子、受信探触子と試験体との距離 y_1, y_2 を試験体中央に集束超音波探触子の焦点を合わせるように配置する。
- ◆しかし、試験体表面での屈折により試験体内の焦点がずれる上、球面収差で焦点が一意に決まらない。
- ◆さらに試験体内部の多重反射で受信信号の振幅変動が発生する。
- ◆図2の亚克力試験体を曲率半径 40[mm]、直径 20[mm]の集束超音波探触子で $y_1:18$ [mm]、 $y_2:30$ [mm](条件1)、 $y_1,y_2:35.5$ [mm](条件2)で透過信号分布を画像化した。
- ◆図3、図4より条件1では $d:1-4$ [mm]の隙間が画像化され、条件2より 14.1[dB]高感度だった。条件2は $d:1-3$ [mm]は画像化できなかった。
- ◆送信時は屈折、球面収差を考慮して近い探触子配置、受信は受信信号の振幅変動を考慮して送信時より遠い探触子配置が必要となる。

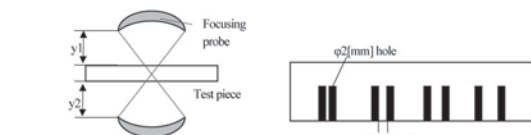


Fig.1 General focusing probe setting.

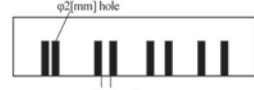


Fig.2 PMMA block for evaluating axial resolution.



Fig.3 Transmission signal intensity distribution at $y_1:18$ [mm] and $y_2:30$ [mm].



Fig.4 Transmission signal intensity distribution at y_1 and $y_2:35.5$ [mm].

1-8-9

1-8-9 Wave mixing 法における Lamb 波の 2 次高調波発生とき裂検出の一検討

A study of second harmonic Lamb waves using wave mixing technique for crack detection

○福田誠, 西平守正, △佐久間裕太(秋田大)

- ◆Lamb 波に wave mixing 法を適用した場合において、き裂位置を固定して wave mixing の位置を変えたときの、2 次高調波成分の検出量の分布について、有限要素解析により検討した。
- ◆長さ 100 mm、厚さ 1 mm の平板中に、 $x = 30$ mm の位置に閉じたき裂を配置し、平板表面の y 方向の振動変位に含まれる 2 次高調波検出量から、き裂位置の特定について検討した。
- ◆一方向からの送波による 2 次高調波検出(従来法)と比較して、wave mixing 法では 2 次高調波の検出範囲がき裂付近に集中すること、き裂から離れた位置では 2 次高調波検出量がより小さくなることから、き裂位置の特定に有効である可能性が示された。

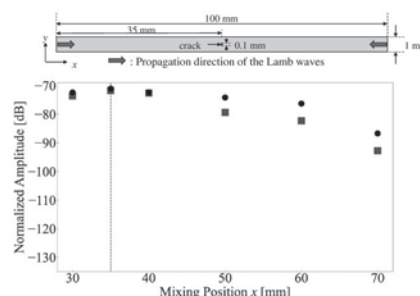


Fig.1: Simulation model and second harmonic components at each mixing position. (●: conventional method, ■: wave mixing technique.)

1-8-10

1-8-10 居室内の単一超音波センサによる見守りの実現

Development of an indoor human monitoring system using a single ultrasonic sensor.

○今村俊樹, △福土和義, △池田賢二(セコム)

- ◆セコムでは独居者の見守りサービス高度化のため、超音波を使った、人の位置や状態を把握する研究を進めている。
- ◆従来、超音波で広い空間をセンシングするにはセンサを複数箇所に設置し、かつ測定対象(人)が機器を所持する必要があった。
- ◆今回、センサ1台のみの設置で、人が機器を所持する必要なく、8畳間程度の空間をセンシングできる技術を確認した。
- ◆センサは180度指向性スピーカと24chアレイマイク、その間に鋸歯状構造の超音波減衰壁を配置する構造により、空間内の3次元情報を計測し、これを基に人の位置や状況把握を可能とした。
- ◆実環境の検証により明らかになった壁からの反射波ノイズの影響を、独自の信号処理にて排除することで、人の位置や姿勢(立位/座位/臥位)や呼吸有無の安定な検出を可能とした。

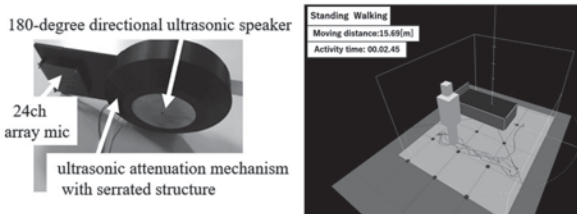


Fig.1: Appearance of ultrasonic sensor

Fig.2: Result of indoor human movement analysis

1-8-12

1-8-12 空中強力超音波源による定在波音場内に濡れた布を入れた時の音圧分布

Sound pressure distribution when a wet cloth is placed in a standing wave sound field by aerial intense ultrasonic sources

☆大和田知花, 浅見拓哉, 三浦 光(日大・理工)

- ◆筆者らは、空中強力超音波源を用いた濡れた布の乾燥について検討を行っている。本稿では乾燥させる試料を音圧の腹または節の位置に入れた時の音圧分布について検討を行った。
- ◆2台の空中超音波源の振動板面が平行かつ向かい合うように設置し、板間に強力な定在波音場を形成した。定在波音場内の音圧の腹または節の位置に乾燥させる試料として濡れた綿布カナキン3号(大きさ44×84×0.5mm)を置き、その設置位置による音圧分布を検討する。
- ◆Fig.1は乾燥させる試料を音圧の節位置に振動板と平行に入れた時の音圧分布の結果である。試料の含水率が高い時は試料付近が音圧の腹になっているが、含水率が低くなるにつれて試料付近が音圧の節になっていることがわかる。

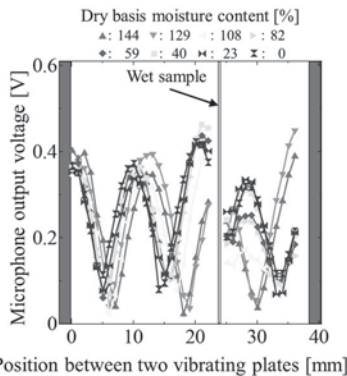


Fig. 1. Sound pressure distribution when a wet sample is placed at node position of sound pressure.

1-8-11

1-8-11 測定光学系の簡素化

電流変調半導体レーザーによる自己遅延ヘテロダイン干渉計を用いた空中音場計測(2)

○中村健太郎, △内山拓己, 和田有司(東工大)

- ◆音場に伴う空気の光学屈折率の変調を光学干渉計で検出する空中超音波測定法がある。音場を乱さない測定が可能である。
- ◆この構成の簡素化・低コスト化をめざして安価な半導体レーザーの電流変調による自己遅延ヘテロダイン干渉を利用した。音響光学変調器を省いた構成で空中超音波の測定ができる。
- ◆本報告では光学系のさらなる簡素化を行った。
- ◆2つのハーフミラー間の多重反射を利用することで、1光軸構成となり、Foot printを減少した。
- ◆コリメートレンズや受光素子の表面反射の利用により、ハーフミラーも省略できた。
- ◆1/Q検波により超音波による位相変調を検出し、絶対音圧を求めた。実際の音圧に近い値が求められた。

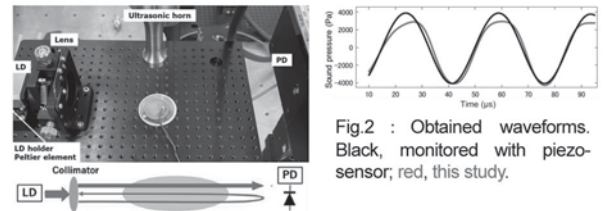


Fig.1: Simplified setup of inline interferometer for ultrasonic field measurement.

Fig.2: Obtained waveforms. Black, monitored with piezosensor; red, this study.

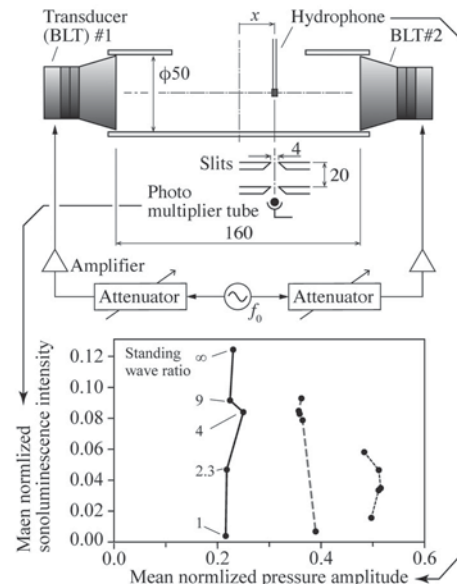
1-8-13

1-8-13 28kHzソノリアクターにおけるソルミネッセンス強度の定在波比依存性

Dependence of sonoluminescence intensity on standing wave ratio in 28 kHz sonoreactor

○黒山 喬允, 小笠原 英子, 森 和義(防衛大)

- 対向したボルト締ランジェバン振動子(BLT)と円筒管からなる定在波比を変更可能なソノリアクターを作成した。
- ソルミネッセンス強度と音圧振幅の分布を計測した。
- 定在波比が1に近づく、すなわち音場が進行波のようになるにつれてソルミネッセンス強度が低下した。



1-8-14

1-18-14 高周波強力超音波照射によって誘起される屈折率変化とナノバブル密度分布の関係

Refractive index change and nanobubble density distribution induced by high-frequency, high-intensity ultrasonic irradiation

☆原田裕生(同志社大), 石河睦生(桐蔭横浜大), 松川真美, 小山大介(同志社大)

- ◆水中への高周波強力超音波照射によって誘起される屈折率変化とナノバブル密度分布(体積分率分布)の関係性について評価した。
- ◆光伝搬シミュレーションを通して導出した屈折率勾配モデル(2次元ガウス分布)および有効媒質近似の Bruggeman モデルを使用し、超音波トランスデューサ表面付近の局所的な位置におけるナノバブル体積分率を算出した。
- ◆超音波トランスデューサ表面付近にて、屈折率変化が最大となる位置($\Delta n = 0.04$: フィッティング値)での体積分率は0.12であった(Fig. 1)。

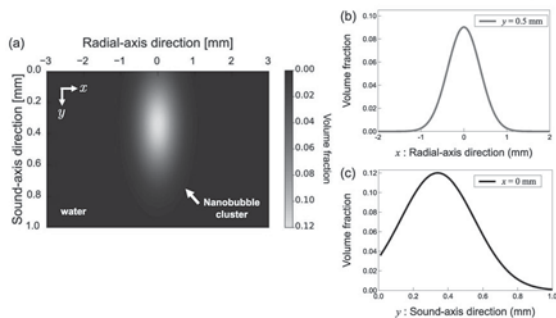


Fig.1 (a) Profile of the volume fraction of the nanobubble cluster near the surface of ultrasonic transducer calculated by an effective medium approximation (EMA). (b) The change in the volume fraction in the sound-axis direction at $x = 0$ mm. (c) The change in the volume fraction in the radial-axis direction at $y = 0.5$ mm.

1-8-16

1-8-16 ダマングレーティングへの位相シグネチャーの応用

Application of phase signatures to dammann grating

○塚安祐輔(筑波大院), 伏見龍樹(筑波大)

- ◆超音波フェーズドアレイにより複数焦点を生成するには通常、複雑な計算が必要である。
- ◆筆者らは前報で、ダマングレーティングを用いて単純な計算で複数の焦点を生成する手法を報告した。
- ◆空中超音波の分野では、ツイン・ボルテックストラップと呼ばれる位相シグネチャーを加えて、特徴的な圧力分布を出力する手法がある。
- ◆本報では、ダマングレーティングに位相シグネチャーを加えて、複数のツイン・ボルテックストラップが生成されることを検証した。
- ◆複製されたトラップでは、中心へと引き寄せる力場となり、トラップとしての性質が保たれていることが分かった。

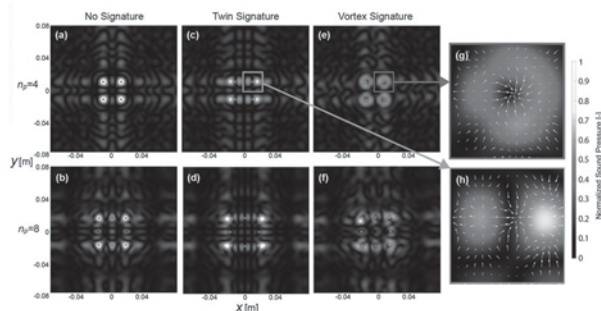


Fig.1: Sound pressure and acoustic radiation force profile on focal plane ($z=100$ mm).

1-8-15

1-8-15 レベルセット法最適化における初期形状と観測点数依存性

—超音波振動工具のトポロジー最適化(6)—
Dependence on the initial holes and observation points on level-set method optimization

—Topology optimization of large ultrasonic tools (6)—

○和田有司, 中村健太郎(東工大・未来研)

- ◆超音波接合等を目的とする振動工具の作用面を平坦に振動させることは産業応用上重要であるが、経験に基づくノウハウとなっている。
- ◆著者らは商用有限要素法プログラムを利用したレベルセット法トポロジー最適化を用いて振動工具の設計方法を検討している。
- ◆複数の初期形状を検討した場合の最適化結果と、Fig.1 のように作製した工具作用面のLDVによる振動測定結果の比較検討を行った。
- ◆Fig. 2 に試作した工具の平坦度(v_{min}/v_{max})について計算結果と測定結果の相関を示す。両者は近い相関関係を示しており、作製した工具は最適化計算結果をよく反映していると考えられる。

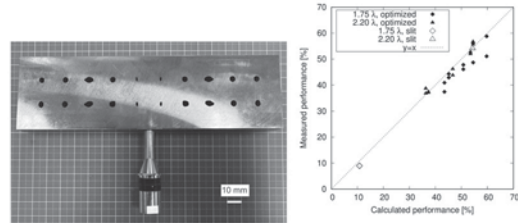


Fig.1: Photograph of the optimized ultrasonic tool.

Fig.2: Correlation distribution between the optimized calculated value and the measurement result for uniformity of the vibration distribution.

1-8-17

1-8-17 非線形集束超音波の各高調波成分の位相補償を利用した鋸歯状波の再構成*

Reconstruction of sawtooth waves using phase compensation of each harmonic component of nonlinear focused ultrasound

◎清水 鏡介(愛媛大院) 大隅 歩(日大理工) 伊藤 洋一(日大理工)

- ◆非線形チャープ信号とミキシングによる周波数補間を用いたパルス圧縮方法について研究を行っている。
- ◆本報告では、ミキシングに対する位相シフトの影響を補償するために、非線形集束超音波の各高調波成分の位相補償を利用した鋸歯状波の再構成を行い、その有効性について検証した。
- ◆Figure 1 に提案手法の概要図を、Figure 2 に実験結果を示す。
- ◆その結果、提案手法の適用により、鋸歯状波が再構成されている様子が見え、提案手法の有効性を確認した。

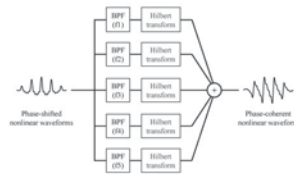


Fig. 1 Schematics of phase compensation of each harmonic component of nonlinear focused ultrasound.

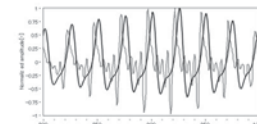


Fig. 2 Experimental result.

1-8-18

1-8-18 強力空中超音波の非線形性を利用したパルスインバージョン法による高調波イメージング
 Harmonic Imaging by Pulse Inversion Method Using High Intensity Airborne Ultrasound with Nonlinearity

☆神谷大樹(日大院・理工), 清水鏡介(愛媛大院), 伊藤洋一, 大隅歩(日大・理工)

- ◆強力空中超音波の非線形性を利用した高調波イメージング(Harmonic Imaging: HI)を提案し, 研究を行っている.
- ◆空中超音波励起による HI にパルスインバージョン法(Pulse Inversion : PI)を組み合わせて, 金属製薄板内欠陥の HI を行った.
- ◆その結果, PI 処理をすることで欠陥位置において反射や回折現象がより鮮明に表れ, 高調波イメージングが行えることを確認した.

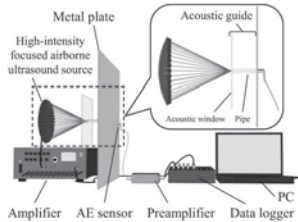
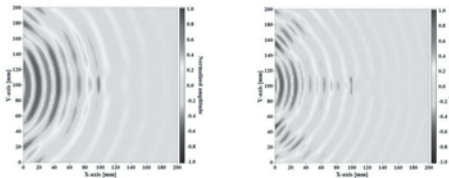


Fig.1 Experiment devices.



(a) Results before applying the PI. (b) Results of applying the PI.

Fig.2 Guide wave propagation image.

1-8-20

1-8-20 講衛星放送用パラボラを用いた聴覚閾値計測の試み

-VHF 領域の聴覚閾値系測定システムの構築(2)-

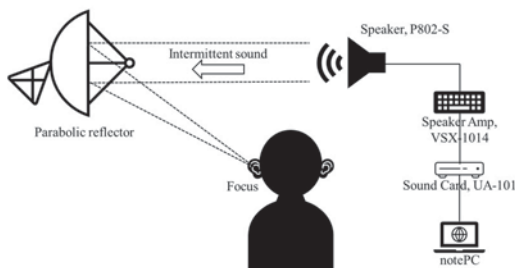
Focusing of VHF sound using parabolic reflector for receiving satellite broadcast, Measurement of hearing threshold at VHF and instrumentation (2)

☆金子武, 大石まなか(近藤奏海(神奈川工科大),

廣江正明(小林理研), 中村健太郎,

Zuhal Maulana Firdaus(東工大), 上田麻理(神奈川工科大)

本研究では VHF 領域の聴覚閾値測定のための検査システムの構築を目的として, その第一段階として VHF 領域の検査音が歪みがなく高音圧で曝露であるパラボラアンテナシステム構築と検査の可能性を探った.



Block diagram of measuring system with parabolic reflector and a participant.

1-8-19

1-8-19 衛星放送用パラボラを用いた VHF 音の集束

VHF 領域の聴覚閾値計測と測定システムの構築(1)

☆Zuhal Maulana FIRDAUS, 和田 有司(東工大), 中村 健太郎(東工大), 上田 麻理(神奈川工大)

- ◆ For the purpose of study of human hearing threshold, a very high frequency (VHF) sound source using an offset parabolic reflector was fabricated.
- ◆ A combination of offset parabolic reflector for BS antenna and tweeter achieved the sound pressure level (SPL) above 120 dB from 15 to 25 kHz.
- ◆ Diameter of the reflector was 560 mm and the focal length was 330 mm.
- ◆ The frequency characteristics were compensated to achieve quasi-flat frequency responses by dynamically adjusting the voltage applied to the tweeter.
- ◆ A mean SPL of 125.4 dB was achieved using the reflector.

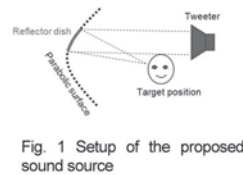


Fig. 1 Setup of the proposed sound source

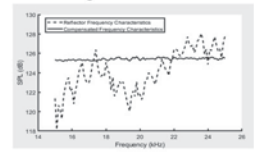


Fig. 2 Obtained SPL between 15 to 15 kHz.

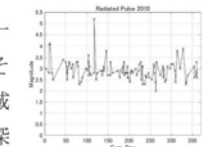
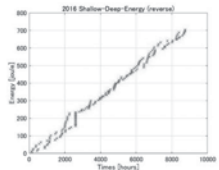
1-8-21

1-8-21 紀伊半島南東沖地震の一次発震元とその発震特性

First hypocenter and its characteristics of the earthquake southeast off the Kii Peninsula.

○菊池年晃(防衛大)

中部海域には地震を頻発する震源域が幾つか存在する。その中で最も活動的な紀伊半島南東沖震源域は, 地域的に限定されているだけでなく深度的にも二分化されている。それらは深度 50km 付近の浅部震源域と深度 400km 付近の深部震源域とが存在する。2016 年に, これらの震源域で発生した累積エネルギーを上図に示す。深部の累積エネルギーを+で, 浅部の累積エネルギーを+で示した。両震源域でのエネルギー推移は一致して, 浅部域のエネルギーは深部域のそれより時間的な遅れがある。即ち, 深部域での発震が重要である。そこで 2010 年に深部域で発生した地震のマグニチュード M を下図に示した。平均的な値 3 の中で, M 5.3 のパルス的な発震が表れた。この発震分布は観測データ(2004 から 2023 年)のほぼ全般で見られた。そのパルスの発震は 4 月から 8 月の範囲に発生している。即ち夏至の前後に発生していることから月の引力が発震に寄与していると考えられる。従って, 地球のなりわいに関連した発震であることから, Statistical Energy Analysis(SEA)では基本発震源の音源と見なせる。従って, 地震の SEA 的伝搬モデルが構築できる。それを元に地震予知モデルを構築できる。



1-8-22

1-8-22 冬季における日本海(Sea of Japan)の音波伝搬

Sound propagation of NIHONKAI (Sea of Japan) in winter

○鶴ヶ谷芳昭 (三陽精工), 菊池年晃 (防衛大)

- ◆日本海と外洋とは水深約 200 m の海峡でつながっている。冬季の冷却により、密度を増した海水は日本海の底部に沈み、冷たい日本海固有水を形成する。秋季に比して、冬季は海面のさらなる冷却によって海面付近の水中音速は低下すると共にサウンドチャンネル(SC)軸深度は上昇する。一方で日本側の大陸棚を出た領域では余剰深度を生じ、収束帯(CZ)伝搬を生じる。
- ◆Fig. 1 に冬季における検討ラインの双方向の音波伝搬を示す。a)は大陸側(NW side)から能登半島側(SE side), b)は能登半島側から大陸側への音波伝搬である。海面付近の音速が低下し、SC 軸深度が上昇したことにより、大陸側ではSC 伝搬を形成している。一方、能登半島側では、余剰深度を生じ、CZ 伝搬が明確になる。また、秋季の大陸側において、海面ダクト伝搬を生じたが、冬季、海面付近の冷却によって音速が低下、SC 軸の上昇によってSC 伝搬が形成される。

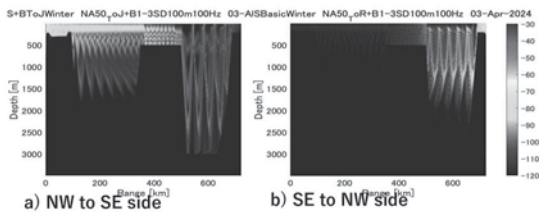


Fig. 1 Comparison of sound propagation for bi-directional propagation in winter season SD;100m

1-8-24

1-8-24 鳥島近海の地震に伴う海中伝搬波の解析

Analysis of underwater acoustic waves associated with earthquakes near Torishima Island

○岩瀬良一(JAMSTEC)

- ◆2023年10月9日の午前4時台から6時台(JST)にかけて鳥島近海で発生した地震に伴う極めて顕著な海中伝搬波が、海底に設置されたS-net観測点のみならず、一部の陸上観測点の地震計でも観測された。
- ◆前報[1]に続きこの海中伝搬波の解析により得られた知見について報告する。
- ◆一連の直達海中伝搬波は同日の06:20過ぎを最後に検出されていないが、その後も約1時間にわたり、後続波が確認できる。これらは観測網から見て音源とは別の方位から到来した海中伝搬波と考えられる。
- ◆そこで、直達波による擾乱がない7時台の波形について、この音源の推定を行った結果、音源の位置は天皇海山群と推定された。つまり、天皇海山群からの反射波と推定される。この推定音源と直達海中伝搬波の音源及び観測網(黒点)との間の音波の伝搬時間はどちらもおよそ30分であり、両者を足し合わせた往復伝搬時間の1時間は、直達波の後1時間にわたって後続波が観測された事実と整合している。
- ◆各波形には直達波に加え、天皇海山群とは別の反射波と思われる波形も後続しており、今後これらについても詳細に検討する。
- ◆また、前報報告した陸上観測点とS-net観測点の波形における卓越周波数の相違は機器特性の相違に起因している可能性がある。

参考文献

[1] 岩瀬, 音講論(春), 57-58, 2024..

1-8-23

1-8-23 エルミート補間型スペクトル法を導入した弾性放物型方程式法による海洋の地震-音響連成解析

Coupled seismic-acoustic analysis of the ocean using elastic parabolic equation method with Hermite-interpolated spectral method

☆川崎圭一郎, 鮫島俊哉(九大・芸工)

- ◆波動理論に基づいた"Range dependent"な解析手法である放物型方程式法にスペクトル選点法を導入する定式化を行った。
- ◆この2手法を組み合わせることで大規模な海洋音場の解析において計算負荷の低減が期待される。
- ◆エルミート補完型スペクトル法を用いることで、1階微分を含む境界条件を近似せずに導入できる。
- ◆本研究で行った定式化では、妥当な計算が行えたものの、従来手法と比較して計算精度が悪くなる可能性が示唆された。

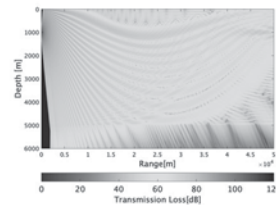


Fig. 1 Proposed Method

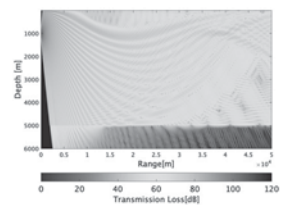


Fig.2 Conventional Method

1-8-25

1-8-25 浮体式洋上風力発電稼働時に発生する水中雑音の伝搬シミュレーション

Simulation of underwater sound wave propagation of noise generated during Operating offshore wind power generation

☆小松史弥(神奈川大院・工学研), 森淳一, 土屋健伸, 遠藤信行(神奈川大・工)

- ◆洋上風力発電稼働時に発生する騒音問題がある。そこで洋上風力発電稼働時に環境負荷を与えるかどうかアセスメントとして調べる必要がある。
- ◆本稿では、海中音の計測手法・評価手法のガイダンスに基づき、「沿岸方向」、「沖合方向」、「等深方向」の3通りにおける、五島列島沖の海底地形を再現し、浮体式洋上風力発電稼働時に発生する水中騒音の伝搬シミュレーションを行った。
- ◆Fig.1には、音源中心深度である38[m]の各海底地形における、騒音の伝搬距離-伝搬損失特性を示す。その結果、深度変化によって伝搬特性に3dB程度の差があることが分かった。

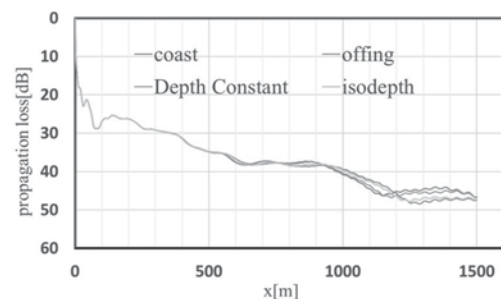


Fig.1: Characteristics of propagation distance and loss (y=38 [m])

1-8-26

1-8-26 ループ管型熱音響プライムムーバーにおけるアルミナボールスタックの実験的検討 —アルミナボールをプライムムーバースタックとした発振実験—

Loop tube type thermoacoustic prime mover experimental study of alumina ball stacks-Oscillation experiment using alumina balls as prime mover stacks-

☆中澤俊貴, 坂本眞一, △平山智士 (滋賀県立大)

- ◆ ループ管型熱音響プライムムーバーを用いたアルミナボールスタックの発振温度に関する検討を行った。
- ◆ ハニカムセラミックススタックとアルミナボールスタックの発振温度を測定し比較を行った。
- ◆ ハニカムセラミックススタックとアルミナボールスタックの発振温度は、525 K と 720 K となった。(Fig. 3, Fig. 4)

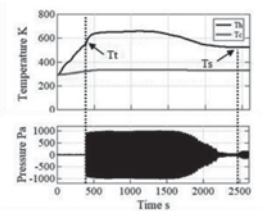


Fig. 3 Time variations of honeycomb ceramics stack temperature and acoustic pressure amplitude.

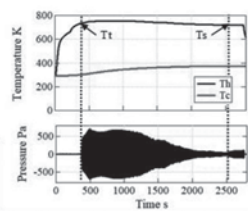


Fig. 4 Time variations of alumina balls stack temperature and acoustic pressure amplitude.

1-9-2

1-9-2 音響メタサーフェス型巾木の吸音性能に関する基礎的検討

Sound Absorption Performance of Acoustic Metasurface Type Baseboards.

☆竹間一心(近畿大院), 菅原彬子, 平栗靖浩(近畿大),

△梶村康平, △武田直樹, △植松成美(東リ)

- ◆ 廊下で発生する歩行音を低減するため、スリット型共鳴器を用いた音響メタサーフェス (AMS) 型巾木を設計し、吸音性能を検討する。
- ◆ Case1, 2は6つの共鳴器を並列させたモデルであり、Case2では斜めネックを用いて吸音性能を同等以上に保ちつつ薄型化を図った。
- ◆ Fig. 1に示す有限要素法解析の結果より、どちらのモデルも歩行音で支配的な周波数付近を吸音していることが確認された。
- ◆ Case2の吸音性能がCase1と概ね一致し、斜めネックを用いた薄型化を実現した。ただし、単一の時に吸音率が最大となる角度がAMS型にした場合でも高い吸音率を示すとは限らないことが分かった。

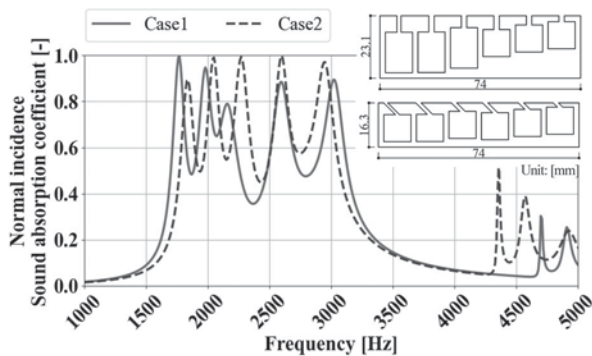


Fig. 1: Results of numerical analysis: Case1 and 2.

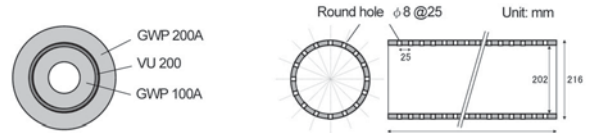
1-9-1

1-9-1 有孔管とグラスウールによる組み合わせ構造の吸音特性に関する実験的検討

Experimental study on the absorption characteristics of the combination structure of perforated pipes and glass wool

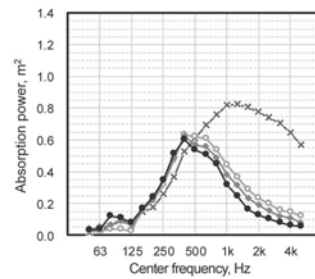
○小泉 穂高 (戸田建設)

有孔加工したVU管とグラスウールパイプ (GWP) を組み合わせた立体吸音体について、残響室法による吸音力測定を行った。VU管とGWPの組み合わせ条件やVU管端部開口の塞ぎの有無によって、GWP単体における吸音力を上回る場合、下回る場合が様々な観測された。



Combination of the perforated VU pipe and glass wool pipes

Perforation condition of the VU pipe



Absorption power of the combination structure of perforated VU 200 and GWP 100A

1-9-3

ホール・劇場椅子の吸音測定における課題とP/A法の適用

Measurement of Sound Absorption of Auditorium Seats and Application of the P/A Method

○西原法子(竹中技研)

客席椅子ブロックの吸音は、無限大試料の吸音率 α_{∞} を得ることを目的として、衝立で周囲を覆って測定するが⁽¹⁾(JIS法)、本稿ではBradleyが提案する客席ブロックの周長(Perimeter)と面積(Area)の比と有限大試料の吸音率 α_s の線形性の仮定を用いた周辺効果の補正方法(P/A法)によって α_{∞} を導出し、2種の実寸椅子で検証した。

測定の結果、吸音性が低い椅子では、 $\alpha_s > \alpha_{\infty}$ (JIS)とならないこと、またP/A法で求めた無限大試料の吸音率とは $\alpha_{\infty}(P/A) \neq \alpha_{\infty}(JIS)$ となることが示された。

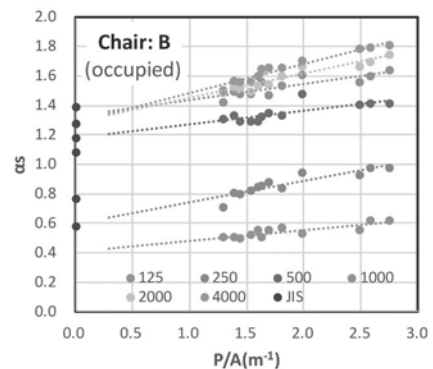


Fig. 3 α_s versus P/A ratio in occupied condition.

1-9-4

1-9-4 コンパクト建築設計資料集の 教育施設諸室の規模検討 —吸音設計の基礎資料として—

Study of the scale of educational facilities based on Compact architectural design data collection. Fundamentals of Absorption Design.

○西川嘉雄(徳山高専)

- ◆吸音設計の基礎データとしてコンパクト建築設計資料集の教育施設のデータから、実在する諸室の規模の検討を行った。
- ◆設計資料集成に記載されている41物件の図面から513室(天井高既知:223室、天井高不明290室)の概略寸法を求め検討した。
- ◆短辺:長辺比は概ね1:1~1:2の間に分布していた。1:3~1:4はワークスペースやオープンスペース、職員室などで細長い空間であった。1:4以上はホワイエ、エントランス、廊下であった。
- ◆天井高は2.5~6.6mである。3m以下の室が73%で、3.5m以下では92%であった。
- ◆今後これらの資料を残響時間の予測や必要吸音率の検討に活用する。

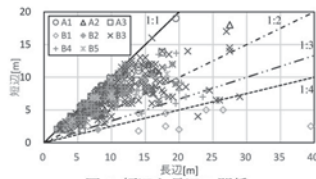


図-1 短辺と長辺の関係

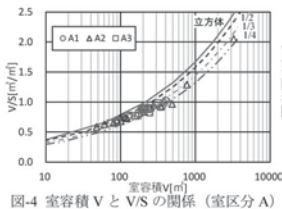


図-4 室容積VとV/Sの関係(室区分A)

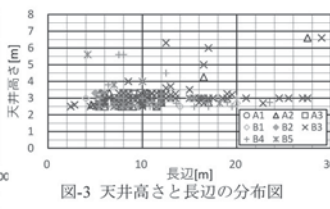


図-3 天井高さと長辺の分布図

1-9-6

1-9-6 物理法則に基づくニューラルネットワーク を用いた音響シミュレーション

Acoustic Simulation using Physics-Informed Neural Networks

◎任逸, △山本篤史(ソニー)

- ◆近年、Physics-Informed Neural Networks (PINNs)を用いた物理シミュレーションが発展し、室内音響におけるアプリケーションも見込まれている。
- ◆従来のデータドリブ学習に対し、PINNsの最大の特徴は深層学習ライブラリの自動微分機能を利用し、物理支配方程式である偏微分方程式を損失関数とし、この損失関数が最小となる学習を行うことである。この物理的制約を課することで、従来のデータドリブ学習より高精度な結果が得られる。
- ◆本稿では、PINNsを用いた3次元大空間音響シミュレーションおよびFDTDなどの教師データが不要な学習例を紹介する。

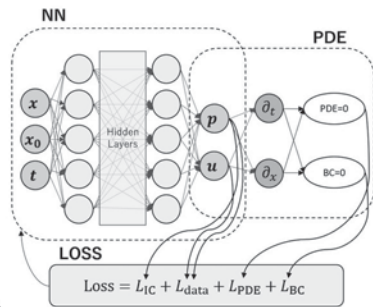


Fig.1: An example PINN model for acoustic simulation

1-9-5

1-9-5 アンサンブル平均を用いた吸音特性の現場測定法 -PU センサの校正に関する再検討-

An in-situ measurement method using ensemble averaging technique for sound absorption of materials

- Further investigations on PU sensor calibration -

○大鶴 徹, 富来礼次, 岡本則子(大分大・理工)

- ◆アンサンブル平均を用いる材の吸音特性測定法(EA_{PU}法)を残響室で繰り返し実施しPUセンサの感度の変化を記録するとともにグラスウールの吸音率を測定した
- ◆音圧、粒子速度ともに感度が変化することを確認した(Fig.1)
- ◆現場校正により感度の変動は抑えられ、互いに近似の吸音率が得られることを確認した
- ◆800 Hz~1 kHzとその倍音の周波数域で生じる感度の微細な上下変動が吸音率へ与える影響に関し、単純な平滑化処理で抑制可能であることを示した

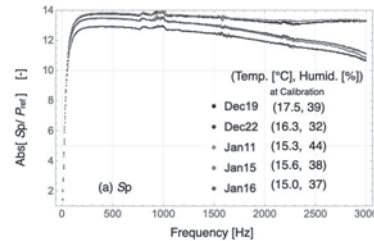


Fig.1: Comparisons of absolute values of relative sensitivities to reference microphone between measurement days. (a): sound pressure Sp, (b): particle-velocity Su.

1-9-7

1-9-7 室内音場の伝達関数における 周波数変動係数の変動要因

Factors affecting the frequency domain coefficient of variation in the transfer function of a room sound field

◎鈴木諒一, 星和磨, 羽入敏樹(日大・短大)

- ◆縮退の影響を定量化する手法として、モード密度や分布の一意性に着目した指標が提案され、スタジオ等の音響設計に用いられている。しかし、それらの指標は固有周波数の正確な把握が必要なため、直方体室への適用のみに留まり、実音場において縮退の影響を評価する測定可能な物理量は確立していないのが実状である。
- ◆実音場で縮退の影響を評価可能にするための物理量として、周波数領域におけるエネルギーのばらつきを評価する周波数変動係数(FCV)に着目している。
- ◆FCVを新たに定式化しその変動要因について検討した結果、周波数離散化幅、モード密度、モードのエネルギーのばらつきの3つのパラメータによってFCVが変動することが分かった。
- ◆周波数離散化幅に依存する問題を解決するためFCVを再定義した。
- ◆室寸法比の異なる音場におけるFCVとモード密度算出した結果、モード密度とFCVに相関があることを示した。

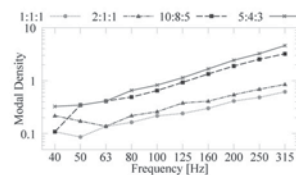


図1 各音場のモード密度

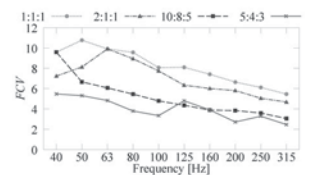


図2 各音場のFCV

1-9-8

1-9-8 壁面の吸音率分布が室内固有モード連成効果に及ぼす影響に関する考察

Consideration of influence of wall absorption distribution on acoustic mode coupling.
☆花山諒平, 高野靖, 大谷真(京大院・工学研)

- ◆閉空間の音響固有モードは互いに独立であるとして扱われることが多いが、境界面の吸音率が大きい場合はモードが連成することが知られている。
- ◆FDTD 法により解析した矩形室(Fig. 1)の音場をモード分解することで、境界面の吸音率分布の違いによる固有モードの連成現象の発生や、音場の減衰過程におけるモード連成の影響について考察した。
- ◆境界面の一部分が他の境界面に比べ大きな吸音率をもつ場合、tangential mode や oblique mode の一部では、モード連成により減衰率が理論値に比べ小さくなることが確認された(Fig. 2)。

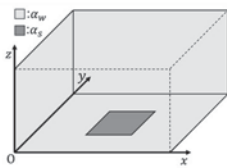


Fig. 1: Calculated rectangular room with absorption plate.

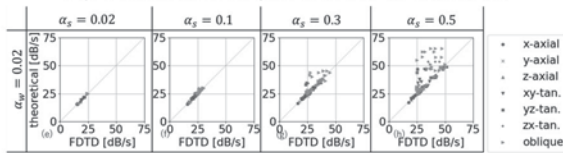


Fig.2: Comparison of theoretical(vertical axis) and calculated(horizontal axis) modal decay rate.

1-9-10

1-9-10 A field survey on acoustic privacy for open plan offices: acoustic and noise measurement

○Ainun Nadiroh, Naoya Maruyama, Keiji Kawai (Kumamoto University)

- ◆Considering the advantage of open-plan offices (OPOs) in gathering many people in a small area, this room may be used in offices to promote interaction among employees. However, information can easily leak out of this room, meaning staff may lack privacy within the room.
- ◆This study describes the acoustics condition of an OPOs. Acoustic measurements were performed to obtain ISO 3382-2 and ISO 3382-3 index values during non-operating hours. In addition, noise level measurements were measured at four points in the office during business hours, and noise events were observed. Correlation analysis was tried to find how the SPL correlations between measurement points describe acoustics privacy.
- ◆Our results show that the office has met the ISO 3328-2 index value standard, but with regard to the ISO 3328-3 standard, the results show a lack of privacy in the office. In addition, the correlation analysis can describe the closeness of acoustic conditions between two locations, however, when it was confirmed by noise event analysis, only a few events related to speech privacy were found.

1-9-9

1-9-9 Barron の修正理論を用いた音声明瞭度指標に関する考察

Study on speech intelligibility indices based on Barron's revised theory
☆陳科吉, 佐久間哲哉(東大・工)

- ◆室内音場における拡散音エネルギーの距離減衰を考慮した Barron の修正理論に基づき、 C_{50} 及び U_{50} の理論値への影響を室条件のパラメトリックスタディにより考察した。
- ◆吸音率が高くなると、 C_{50} の補正值は受音点が平均自由行路より近い場合に負、遠い場合に正となることが確認された。
- ◆一方、 U_{50} の補正值は近傍点だけでなく遠方点でも負となるが、10m 付近の受音点ではほとんど影響がないことが明らかとなった。
- ◆中・大規模講義室の実測値との比較検討では、Barron の修正理論により近傍点及び遠方点における U_{50} の理論値が低下し、全体的な対応性が若干向上することも確認できた。

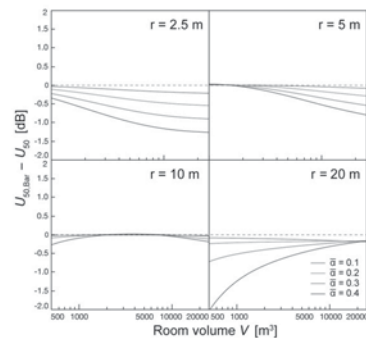


Fig : Difference between $U_{50,Bar}$ and U_{50} for 4 source-to-receiver distances with 4 absorption coefficients.

1-9-11

1-9-11 一般次元 Helmholtz 方程式の極座標解についての考察

Study on the Polar Coordinate Solution of the General-dimensional Helmholtz Equation.

○岩見貴弘, 井上尚久, 尾本章(九大・芸工)

- ◆我々はこれまで一般の空間次元のレイ信号処理手法を提案してきた。このような一般次元での理論構築は、学習効率の向上やコーディングの時間短縮、物理現象の本質の究明の一助になると考えている。
- ◆本稿では、一般次元の斉次 Helmholtz 方程式の極座標解を簡易な流れで導出する。
- ◆また、動径方向関数を球面状帯域制限空間で正規性を保つように定めることで、幾つかの式が単純な形で書き直せることを示す。
- ◆最後に球面状帯域制限空間の正規直交基底を導出し、再生核の性質から加法定理を導出可能であることを確認する。

1-9-12

1-9-12 一般指向性素子間の相反原理

Acoustic Reciprocity between General Directional Elements.

○岩見貴弘, 井上尚久, 尾本章(九大・芸工)

- ◆相反原理は、音源と受信点間の伝達関数が位置の入れ替えに於いて不変であることを主張する原理である。
- ◆我々は既報で空間微分作用素を用いた指向性受信モデルを提案しており、本稿ではこの受信モデルと多重極展開を用いた放射モデルを用いて一般指向性素子間の相反原理を導出する。
- ◆証明において重要となるのは基本解の空間微分に対する性質であり、これを用いて相反原理の幾つかの別の物理的解釈が可能であることを示す (Fig. 1)。

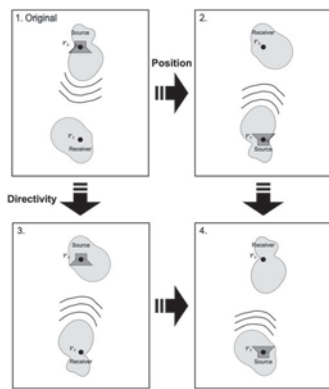


Fig.1 Conceptual diagram of the extended acoustic reciprocity.

1-9-14

1-9-14 仮想球面のコーシーデータに基づく球中心での受信指向性の合成

Synthesis of sound receiving directivities at the center of a sphere based on Cauchy data on the sphere's imaginary surface.

○井上尚久, 岩見貴弘, 尾本章(九大・芸工)

- ◆空間微分により受信指向性を合成する方法が近年提案されているが、空間微分の演算の行い易さは解析手法や音場表現に依存する。
- ◆本研究では仮想球面上のコーシーデータ (音圧・その法線方向微分) から球中心点での受信指向性を合成する手法を構築した。
- ◆八節点六面体要素 (Lagrange-Trilinear 内挿関数) を用いた FEM による解析例を示した。
 - 内挿関数の微分からは算出不能な同一方向の2階微分値についても算出可能であることを示した。

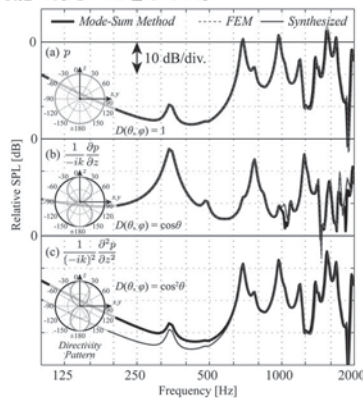


Fig. 1 Level frequency responses: (a) the sound pressure, (b) the first-order derivative with respect to z, and (c) the second-order derivative with respect to z, respectively.

1-9-13

1-9-13 壁面吸音を考慮した矩形室の固有モード展開解析における音源・受信指向性の導入

Introduction of the sound source and receiving directivities in the eigenmode expansion analysis of rectangular rooms with wall sound absorption.

○井上尚久, 岩見貴弘, 尾本章(九大・芸工)

- ◆空間微分により音源・受信指向性を合成する方法が近年提案されており、空間微分の演算の容易な矩形室の固有モード展開解析での解析事例を示した。
- ◆単純な空室の矩形室では計算効率も良好であり、実用的である。
- ◆理論解析解であるため、可聴化や数値解析のためのベンチマーク解の算出方法としての活用も期待される。

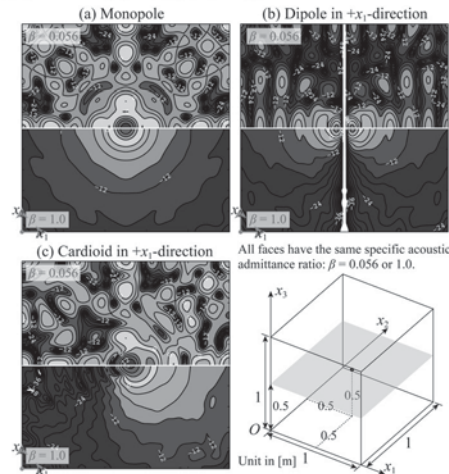


Fig. 1 Relative SPL contours drawn every 3dB for pointwise sources with simple directivities.

1-9-15

1-9-15 FMBEM 入射指向性解析に基づく拡散性指標の空間分布に関する検討

Study on spatial distribution of sound-field diffusion indices based on FMBEM incidence directivity analysis

☆萩原諒(東大・工), 樹本貴之(サイバネット),

安田洋介(神奈川大・建築), 佐久間哲哉(東大・工)

- ◆既報では直方体室や不整形室における定常音場を対象に、FMBEM 入射指向性解析に基づく拡散性指標を計算し、基礎的な検討を行った。
- ◆本報では不整形室内の多点に対して拡散性指標を計算し、吸音条件と指標の空間分布との対応関係について検討を行った。
- ◆種類の拡散性指標 d_θ, d_l について整理し、指向性の打ち切り次数 N_l ごとに一対一の関係で対応することを確認した。(Fig. 1)
- ◆不整形室内の拡散性指標の空間分布では、吸音条件や指標の種類によらず、音源点の付近において指標が著しく低下する様子が見られた。
- ◆吸音面の付近では拡散性が低下する傾向が見られ、床面を全て吸音とした際にはさらにその傾向が顕著に表れていた。
- ◆解析結果から得られた d_θ と d_l の値の違いについては、Fig. 1 の対応関係に従っていることを確認した。

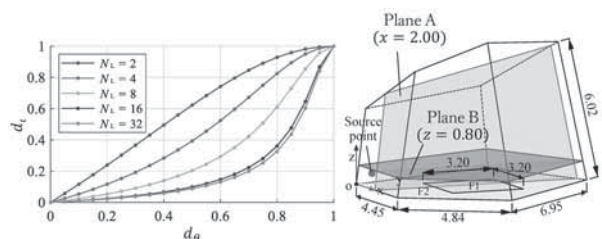


Fig. 1 Relationship between d_θ and d_l . Fig. 2 Analysis model. (unit: m)

1-9-16

1-9-16 複層ガラスの斜入射音響透過損失の数値解析 —中空部の粘性減衰の影響

Numerical analysis of sound transmission loss of double-pane glass under oblique-incident conditions – Viscous damping effect of air space

◎山崎泰知, 田中ひかり, 増田潔(大成建設), 井上尚久(九大・芸工), 佐久間哲哉(東大・工)

- ◆ガラスの遮音性能は、内部減衰や端部の損失の影響を受ける。さらに複層ガラスでは、**中空部での粘性減衰の影響**も受けると考えられる。
- ◆斜入射条件の複層ガラスの音響透過損失の解析で、ガラス中空部の粘性・熱拡散の影響を考慮したモデル化を行い、この影響を検証した。
- ◆Figure 1に示すように、粘性・熱拡散を考慮することで、より実験室実験の結果に近い斜め入射音響透過損失を計算することができた。

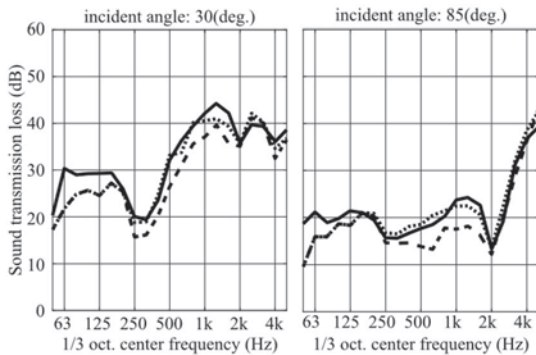


Fig.1: Result of sound transmission loss (solid line: Measurement, dashed line: Numerical analysis w/o thermal-viscous effect, dotted line: Numerical analysis w/ thermal-viscous effect)

1-9-18

1-9-18 複数音源の多点音圧計測に基づく波数領域音響反射率の推定

Wavenumber domain acoustic reflection characteristic estimation based on multi-point sound pressure measurements of multiple sound sources.

☆星加慧, 岩見貴弘, 尾本章(九大芸工)

- ◆筆者らは波数領域の音響反射率を数値計算へ導入することで、任意の材料反射特性をモデル化する手法を提案してきた
- ◆本研究では、材料表面における入射波と反射波の関係から連立方程式を構築し、その逆問題を解くことで波数領域音響反射率を推定する手法を提案する
- ◆2次元音場を対象とした数値計算結果より、提案法で精度良く波数領域音響反射率を推定可能なことを示した

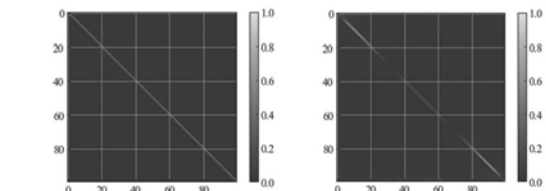


Fig. 5: Estimation result of C_r (specular reflection) Fig. 6: Estimation result of C_r (Bragg reflection)

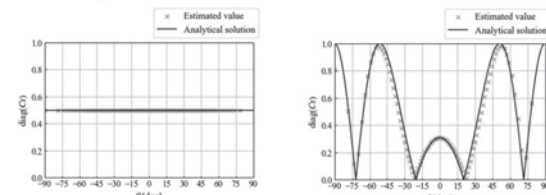


Fig. 7: Diagonal element of C_r (specular reflection) Fig. 8: Diagonal element of C_r (Bragg reflection)

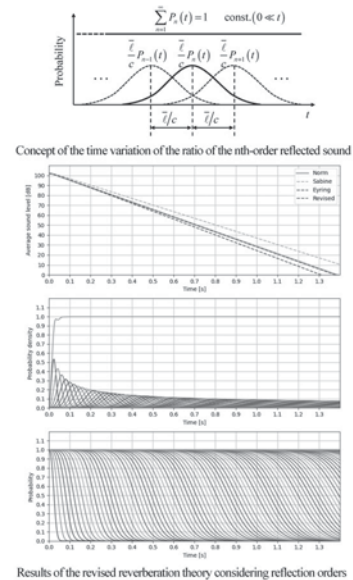
1-9-17

1-9-17 反射次数を考慮した拡散音場における修正残響理論

Revised reverberation theory in a diffuse sound field taking into account reflection orders

○羽入敏樹(日大・短大)

- ◆“直接音の残響”の概念を用いずに、“反射音による残響”として修正残響理論の記述を試みた。反射音の反射次数毎の時間分布に関する数理モデルを提案し、この数理モデルによる残響減衰が修正残響理論および音線法シミュレーション結果と一致することを示す。



Results of the revised reverberation theory considering reflection orders

1-9-19

1-9-19 物理モデルと深層学習を用いた少数マイクロホンによる散乱音場推定

Scattered sound field estimation by a small number of microphones using physics models and deep learning

☆鬼澤諒介, 佐藤元, 津國和泉, 池田雄介(東京電機大)

- ◆複雑な音の伝搬を効率的に理解するには少数の測定情報から音場を推定し、可視化することが有効
- ◆本研究では、少数の計測情報から滑らかな吸音率を持つ物体に対する散乱音場を可視化することを目的として、深層学習に物理モデルを導入した散乱音場推定手法を提案。
- ◆シミュレーション実験では、吸音率を持つ球に対する散乱音場を球形マイクロホンアレイの情報や物理モデルを用いて推定し、その精度を評価。
- ◆実験の結果、高精度に散乱音場を推定可能であることを確認

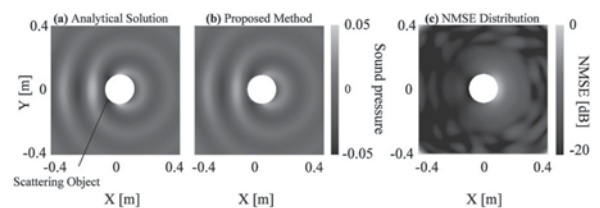


Fig.1: Comparison of scattered sound field using proposed method and analytical solution. (Frequency:2kHz)

1-10-1

1-10-1 改訂騒音障害防止のためのガイドライン
が目指すもの

What is the purpose of noise measurement in the noise guideline revised by the Ministry of Health, Labour and Welfare?

○東久保一朗(産業医大)

厚生労働省は、令和5年4月に、労働現場におけるさらなる騒音障害防止をはかるために、事業者が自主的に講ずることが望ましい騒音障害防止対策の体系として改訂騒音ガイドラインを発出した。改訂点は、対策の実施者としての「管理者」の選任、聴力検査のスクリーニングとして4000 Hzは25 dB, 30 dBで判定する、雇入れ時や二次検査の聴力レベル測定の際に6000 Hzを新たに加える、聴覚保護具の選定および適正な装着を推進するために遮音しすぎない聴覚保護具の選択と装着を義務化、騒音作業時間による管理でばく露低減対策が可能とした。さらに騒音レベルの測定方法として作業環境測定に加えて、新たに個人ばく露測定、最も騒音レベルの高い位置での定点測定、騒音発生源の音響パワーレベルによる騒音レベル推定が追加された。その結果、騒音障害防止対策の基礎となる騒音レベルの測定については、屋内の作業、屋外の作業、手持ち工具作業の様に作業者が騒音発生源となる作業や移動を伴う作業の様に従来の作業環境測定ではカバーできないケースに対して、各々作業者の騒音レベル評価に適切と思われる測定方法により騒音レベルを求めてばく露低減措置につなげることが可能となった。

監督者としての管理者の選任、科学的な騒音測定、より騒音性難聴の傾向を把握するための健康診断、適切な聴覚保護具の装着等のばく露防止対策の体系化は、本邦における労働者の騒音性難聴の予防に役立ち、健康でいきいきとした社会形成に資するものであると確信している。

1-10-3

1-10-3 騒音障害防止ガイドラインの改訂と
日耳鼻産業保健・環境委員会の動向

Revision of the Guidelines for Preventing Noise-Related Disorders and the Action of the Industrial Health and Environment Committee

○和田 哲郎(筑波大)

- ◆騒音障害防止ガイドラインが改訂され、日本耳鼻咽喉科頭頸部外科学会では産業保健・環境委員会を中心に、改訂内容の周知、対応方針の検討、医学的知見の収集に取り組んできた。関連する活動を示す。
- ◆騒音性難聴担当医制度：耳科学と産業医学の両方の専門的知識を獲得するため日耳鼻学会が研修会と試験を行い、騒音性難聴担当医を認定している。担当医は2024年4月現在、全国で1000人を超え、日耳鼻学会ホームページで名簿が公開されている。
- ◆産業保健総合支援センターとの連携：都道府県毎に耳鼻咽喉科医と産業保健総合支援センターの連携状況を調査報告した。
- ◆聴覚管理マニュアルならびに Q&A：産業医や耳鼻咽喉科医が健康診断で適切に聴覚管理を行い、騒音性難聴に関する相談に対応できる様、改訂ガイドラインとも齟齬がないようにマニュアルならびに Q&A を整備し、日耳鼻学会ホームページに掲載した。
- ◆屋外作業場での個人ばく露測定：改訂ガイドラインに沿った個人ばく露測定が可能か、現地調査を実施し、学会で報告した。
- ◆ヘッドホン・イヤホン難聴の啓発活動：ワークグループを立ち上げ、若者のイヤホン等の使用状況調査、難聴防止の啓発活動を開始した。
- ◆Japan Hearing Vision と AC Japan：自民党難聴対策推進議員連盟と協力し、Japan Hearing Vision という提言を社会に向けて発信した。また、日耳鼻学会は AC ジャパンによる難聴啓発キャンペーンを令和6年7月から開始した。

1-10-2

1-10-2 海外事例とともに見る騒音障害防止
ガイドラインの改訂の意義と今後について

The significance of the revision of the guidelines for prevention of noise-induced hearing loss, and its future aspects together with overseas examples.

○柴田 延幸(安衛研)

- ◆約30年ぶりの騒音障害防止ガイドラインの改訂により、我が国の騒音ばく露管理においてばく露許容基準の考え方が導入されるとともに、個人騒音ばく露計の使用が認められるに至った。
- ◆我が国では改訂前の騒音障害防止ガイドラインにおいても85dB以下の騒音作業場を第一管理区分とするなど、主として騒音事業場の管理区分の決定をするために場の測定・管理を実施してきたところであるが、諸外国では個人ばく露測定を軸として騒音作業管理を実施している。
- ◆作業者の一日当たりの騒音ばく露作業が異なる n 種類の騒音ばく露作業である場合、i 番目の騒音ばく露作業の等価騒音レベルを L_{dBA} 、作業時間をも時間とすると、この作業者の作業管理として満たすべき一日当たりの騒音ばく露量は、次式であらわされる。

$$10^{\frac{L_1}{10}} \cdot t_1 + 10^{\frac{L_2}{10}} \cdot t_2 + \dots + 10^{\frac{L_n}{10}} \cdot t_n \leq 10^{\frac{85}{10}} \cdot 8$$

- ◆また、 L_p [dB]: 騒音レベルの推計値、 L_w [dB]: 音響パワーレベル、 r [m]: 騒音源からの距離とすると、屋外作業では騒音作業ごとに次式を用いて騒音源から距離 r の位置における等価騒音レベルの推計を行うことができる。

$$L_p = L_w - 20 \log_{10} r - 8$$

1-10-4

1-10-4 騒音性難聴防止のための
JIS に基づく聴覚保護具の選定方法

Selection method of hearing protection devices based on JIS for preventing noise-induced hearing loss

○横山 栄(小林理研)

- ◆令和5年に騒音障害防止のためのガイドラインが改訂され、騒音性難聴防止に配慮して、等価騒音レベルが85 dBを超える作業場では、管理者は2020年に制定されたJIS T 8161-1,-2:2020(現行JIS)による遮音性能を参考に適切な聴覚保護具を選定し、作業者に正しく着用させることが義務化された。本報では、JISで規定された3種類の聴覚保護具の選定方法について、概説する。
- ◆現行JISはISO規格を対応国際規格として国際整合化が図られ、方法規格として新たに制定された。遮音性能として、125 Hz~8,000 Hzのオクターブバンドごとの遮音値、SNR (single number rating) 値、 H, M, L (high/medium/low-frequency attenuation value) 値が規定され、これらを用いて、それぞれ聴覚保護具着用時の騒音曝露レベル(実効A特性音圧レベル)を推定する3種類の方法も規定された。
- ◆本報では、フォームタイプの耳栓を例に、ある騒音を仮定して、オクターブバンド法、SNR法、HML法により耳栓着用時の騒音曝露レベル(実効A特性音圧レベル)を推定した。その結果、いずれの方法でも70, 71 dBとなり、同等の結果が得られることを確認した。
- ◆なお、実際の作業場では、必ずしもJISによる遮音性能が得られないことも知られており、諸外国では、SNR値などの単一評価指標値に対し、減算値の設定がされている。わが国ではまだ減算値の設定はなく、騒音性難聴防止の観点から、今後の検討が望まれる。

1-10-5

1-10-5 外耳道入口の騒音ばく露測定システムの検討

Development of methods for the determination of the sound exposure from sources located close to the ear

○中市 健志 (リオン)

- ◆航空交通管制官や電話対応作業員など、耳載せ形又は耳覆い形の通信機器の使用によって、作業員本人の耳元付近に騒音源がある場合、場の測定を行う作業環境測定では耳元の騒音レベルを測定することは難しい。
- ◆耳元付近にて通信機器から発せられる騒音を測定するためには、外耳道に収まる程度の小型マイクロホンを用いることで測定可能となる。
- ◆外耳道入口付近にて得られた結果には①作業員自身が発した音声レベル、②通信機器から発せられた通話相手の音声レベル、③職場の環境騒音レベルが含まれているため主成分の騒音を推定する必要があるが、職業上音声を録音してこれらを判別することは現実的でない。
- ◆本報告では、外耳道入口付近におけるマイクロホンで測定するシステムと、机上和室内に設置したマイクロホンにて測定した結果から通信機器から発せられた音声区間のみを推定し等価騒音レベルを算出する試みを紹介する。

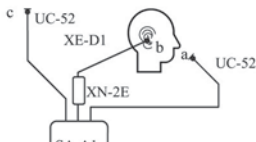


Figure: A system for synchronized recording of signals from three microphones monitoring.

1-10-7

1-10-7 韓国における純音聴力検査機器と環境の適合性調査

Compatibility survey of pure tone audiometry equipment and environment in Korea

○曹 浣豪(KRISS)

- ◆純音聴力評価は聴力検査において最も基本になる方法であり、そのやり方や使用機器の仕様などは国際規格に詳しく定められている。
- ◆しかし、それらの適用に置いて、機器の適合性の検査や校正が国際測定標準に基づいた品質システム下で行われていない場合が多い。
- ◆本研究では韓国の医療現場で使用されている純音聴力検査機器において、国際単位系からトレーサビリティを確保したシステムに基づき、基準との偏差を確認した。
- ◆今後、検査結果のより高い信頼性を保証する体制として、国際測定標準とのトレーサビリティに基づいた品質システムの構築が必要と考えられる。

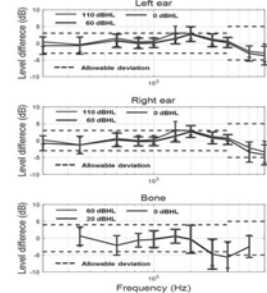
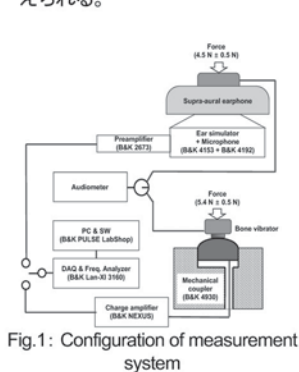


Fig.2: Averaged differences btw. the measured output level and the target level of each indicated output setting

1-10-6

1-10-6 騒音障害防止の啓発活動の事例報告 ~林業関係者に向けた騒音教育~

Case Study on Educational Activities for the Prevention of Noise-Induced Hearing Loss -Education for forestry workers-

◎武田 葵 (リオン)

- ◆騒音障害防止の啓発活動の一環として、林業経営体の事業主に対して、予防に繋がることを目的とした研修を企画し実施した。
- ◆騒音性難聴者の聴力図を模擬した帯域通過フィルタを与えた音声を聴取させ、言葉の聞き取り困難を体験することで、コミュニケーションに障害が起こることを認識させ、予防の重要性を訴えた。
- ◆作業員にチェンソーで丸太を切断する作業を行わせ、B測定位置において100 dB近い騒音レベルに達すること、作業位置周辺の位置において90 dB以上の騒音レベルになることを確認し、どちらの位置でも聴覚保護具を必ず使用する必要があることを認識させた。
- ◆騒音の低減のための対策の例とその効果を実演しながら説明した。発生源対策としてチェンソーの目立てを行うこと、伝ば経路対策として騒音源から離れること、受音者対策として聴覚保護具の使用が効果的であることを理解させた。
- ◆予想以上に騒音レベルが高いということに驚く参加者が多く、普段の作業で騒音対策が求められることの自覚を促すことができた。「イヤーマフの重要性を認識できた」という声が寄せられ、騒音対策への意識が高まるきっかけとなった。
- ◆自身の作業の騒音のレベルを把握させ、それに応じた措置の効果を体験させることは、騒音障害を自分自身に関わる問題であると認識させるよい手法であることが本研修を通じて得られた。

1-10-8

1-10-8 SN比が相互相関関数を用いた音源位置推定精度に及ぼす影響

Effect of signal-to-noise ratio on the accuracy of source location estimation using cross-correlation.

☆森本誠至(近畿大院), 原田和典(岡山県立大), 菅原彬子, 平栗靖浩(近畿大)

- ◆相互相関関数を用いた音源位置推定において、背景騒音と物理的障壁の影響に着目し、これらの要因が推定精度にどのように影響するかを明らかにした。
- ◆背景騒音の影響については、各音源の受音点でのSN比より確認した。背景騒音源の音量が増加すると到来時間差算出に用いる相互相関関数の最大値を取るピークの位置が変化し、背景騒音源の位置が推定された。
- ◆物理的障壁による推定誤差の原因については、回折による伝搬経路長の変化によるものと、受音レベルの低下が引き起こすSN比の低下によるものが確認された。

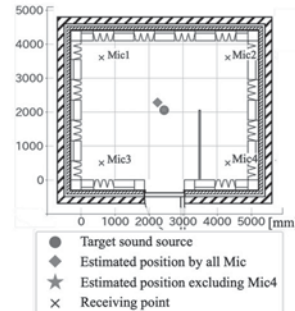


Fig. 1 Estimation error due to diffraction.

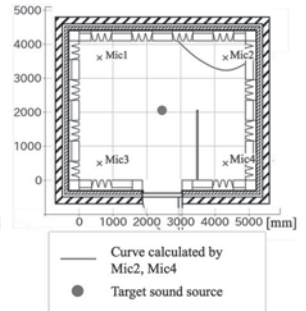


Fig. 2 Estimation error due to reduced signal-to-noise ratio.

1-10-9

1-10-9 ダミーヘッド(HATS)を用いた VHF 領域の耳介周りの音響計測(その8)

—毛髪が VHF 音の伝搬に与える影響—

Acoustic measurement around the pinna in the VHF region using a dummy head—Part8, Effects of hair on VHF sound propagation

☆近藤奏海, 大石まなか(神奈川工科大), 廣江正明(小林理研),

上田麻理(神奈川工科大)

- ◆毛髪による VHF 音(16 kHz~32 kHz の音)の聴こえの影響を調査するため, 先行研究と同様に HATS を用いて, ウィッグなしと 3 種類の長さの異なるウィッグを被せた場合, 4 条件の HRTF を求めた.
- ◆顔正面に音源がある場合を基準(0°)に回転台に置いた HATS を 1° 間隔で一回転させた.
- ◆顔正面からの入射(Fig.1)の場合, 10 kHz 以下では毛髪に依る HRTF への影響は小さい. 毛髪による影響は概ね 10 kHz 超から現れる. また, 毛髪が長いほどその影響(HRTF 間の差)が大きい.

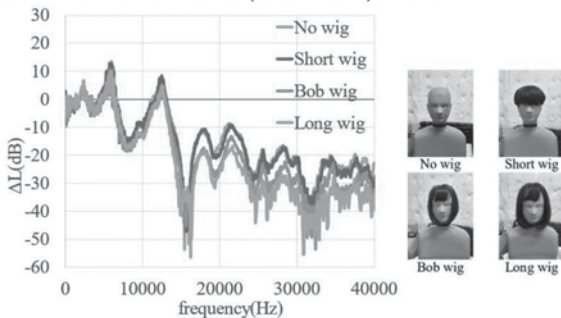


Fig.1 Comparisons of HRTFs among no wig and three wigs at the horizontal angle of 0 degrees

1-10-11

1-10-11 翼騒音制御向け低周波圧電スピーカによる回転音場の生成

Rotating sound field formulation using low-frequency piezoelectric speaker for rotational blade noise control

○高安 俊貴, 後藤 達彦(株)東芝 研究開発センター)

- ◆ファンやモータなどの回転機による騒音, 特に低周波成分の対策において狭帯域の能動騒音制御 (Active Noise Control: ANC) システムが有効とされており, 制御スピーカの小型・軽量化が強く望まれている.
- ◆本報では, 翼騒音制御システムへの応用を目的として, 低周波圧電スピーカによる制御空間への回転音場の生成方法を検討した (Fig.1).
- ◆複数の圧電スピーカユニットを円状に配置し, D 級パワーアンプ基板で構成した圧電スピーカドライバにて 800 Hz の位相差駆動を行った.
- ◆1次・2次ロブモード翼音源特性を模擬できることがわかった(Fig.2).

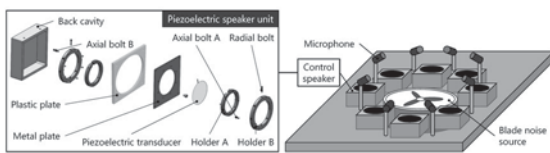


Fig.1:Active noise control system for rotational blade noise.

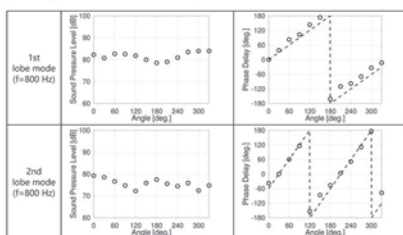


Fig.2:Distribution of sound pressure level and phase delay. (○:Measurement, ---:Ideal value)

1-10-10

1-10-10 螺旋構造を用いた音響レンズのサイレンサーへの適用

Application of acoustic lenses in silencers using Spiral Structures

◎平田武士, 芦澤剛, 宮本光亮, 森尾謙一(日本音響エンジニアリング)

- ◆音響レンズに螺旋構造を採用することで, サイレンサーの3次元形状への拡張に関する検討を行った.
- ◆今回検討した3次元形状の音響レンズの影響を確認したところ, 透過損失は誤差程度であり, 今回調べた周波数範囲では音響的な抵抗になるとは考えにくいことが分かった.
- ◆音響レンズを用い GW を設置すると, 音響レンズを用いず GW を設置したダクトサイレンサーと比べ, 広い帯域にわたり挿入損失が上昇した.

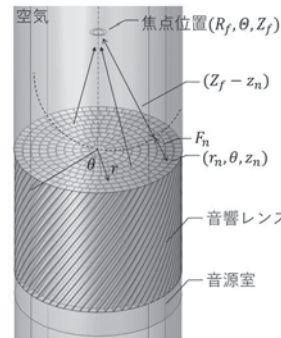


Fig.1:Acoustic lens and silencer with Spiral Structures

1-10-12

1-10-12 風洞試験における自由せん断層内の流速分布に基づく音波の移流量推定

Estimation of sound convection based on a flow profile around a free shear layer in a wind tunnel test

○山崎展博, Δ井上達哉(鉄道総研),

中山雅人(阪産大), 西浦敬信(立命館大)

風洞の開放型測定部で空力音の指向性を測定するうえで, 音波が主流により下流に流される影響や, ノズルまわりに発生する自由せん断層を音波が通過する際に屈折する影響(以下, 移流効果)を考慮する必要がある. 実際の自由せん断層内部では主流に直交する方向に対して流速が連続的に変化しており, またその幅も下流に向かうに従って拡幅しているため, 自由せん断層の流速分布を考慮したうえで移流効果を補正する必要がある.

本研究では, 自由せん断層内部における主流方向の流速成分を実測したうえで, せん断層内を無限小厚さの領域に分割したモデルによる移流効果の補正法を提案する. さらに, 風洞の暗騒音に対して高い信号対雑音比(以下 SNR)が確保でき, 指向性を有するパラメトリックアレイエミッタを活用した評価実験を行い, 提案手法の妥当性を確認した(Fig.1).

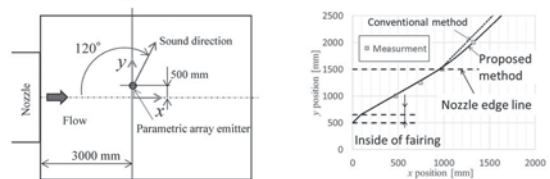


Fig.1 Schematic layout of wind tunnel test(left) and estimated acoustic path advected by the flow(right)

1-10-13

1-10-13 楽器演奏家の音環境に関する研究 (その1) — 個人曝露計を用いた音響曝露調査 —

A study on sound environments of instrumentalists, Part-1 - Survey on sound exposure of instrumentalists using personal dose meter -

○廣江正明(小林理研)、山田翔雅、中島康貴(リオン)、上田麻理(神奈川工科大)

- ◆日常的に大きな音に曝露される職業として、工場や重機、空港駐機場などの高騒音下での作業者が挙げられるが、楽器演奏を専門とする演奏家も大きな音に曝露されている職業の一つである。
- ◆演奏家の適切な音環境や必要な聴覚の保護対策を考えるため、二人の楽器演奏者(サクソ奏者とピアノ奏者)の音曝露の実態を調査した。
- ◆測定には RION NB-14 を用いて1秒毎の $L_{Aeq,1s}$ 、 $L_{Ceq,1s}$ 、 $L_{Cpeak,1s}$ を記録した。
- ◆サクソ奏者がトランペット、ピアノ、ベース、ドラム、歌手、タップダンスら(全員 PA 使用)と、演奏会のリハーサルを行った時の個人曝露計測結果を Fig.1 に示す。
- ◆わずか52分間に1日許容量(85 dB×8時間)の1.2倍相当の音に曝露された事例である(なお、同日に演奏会も行っている)。

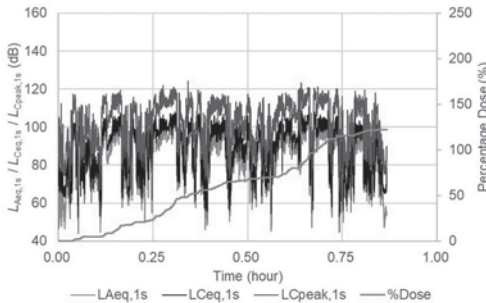


Fig.1: An example of personal sound exposure for an instrumentalist

1-11-2

1-11-2 スペクトログラム無矛盾性に基づく正則化付き独立低ランク行列分析の実験的評価

Experimental Evaluation of Independent Low-Rank Matrix Analysis with Regularization Based on Spectrogram Consistency

☆児島孝明(東大)、高宗典玄(東大)、北村大地(香川高専)、猿渡洋(東大)

- ◆著者らはこれまで分離信号のスペクトログラム無矛盾性に基づく正則化付き独立低ランク行列分析(ILRMA)を提案してきた。
- ◆これまで検証してきた音声信号だけでなく音楽信号で分離実験を行った。
- ◆また、以前の研究で得た観察に基づいて新たに次の手法を追加した。
 1. 最適化の終盤では正則化の重みを減少させるテンパリング手法
 2. Consistent ILRMA と無矛盾正則化 ILRMA を併用する手法 (combination method)

実験結果

- ◆音楽信号の分離においても無矛盾正則化の有効性を確認した。
- ◆また、新たに導入した手法の有効性を確認し、特に combination method が最良の結果を示した。

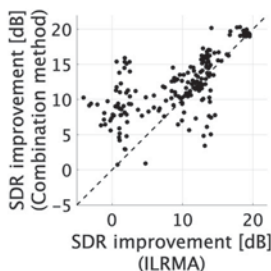


Fig.1: SDR improvements of combination method compared with ILRMA in music signal separation.

1-11-1

1-11-1 空間正則化付き独立低ランク行列分析のための補助関数型ステアリングベクトル推定法

Steering vector estimation in spatially regularized ILRMA based on auxiliary-function method

☆平田爽太、高宗典玄、山岡洗瑛(東大)、北村大地(香川高専)、猿渡洋(東大)、高橋祐、近藤多伸(ヤマハ)

- ◆空間正則化付き独立低ランク行列分析(SR-ILRMA)は、与えられたステアリングベクトル(SV)を正則化項として用いることで、高い分離性能を達成することが知られている。
- ◆本研究では、SR-ILRMAにおいて、分離行列等に加えてSVも同時に最適化することを目指し、majorization-minimization(MM)およびmajorization-equalization(ME)アルゴリズムに基づき、分離行列からSVを推定する手法を提案した。
- ◆独立低ランク行列分析(ILRMA)によって得られた分離行列からSVを推定する実験を行い、収束速度の比較や初期値の検討を行った。
- ◆分離実験により、推定したSVを用いたSR-ILRMAが通常のILRMAと比較して高い性能を示すことを確認した。

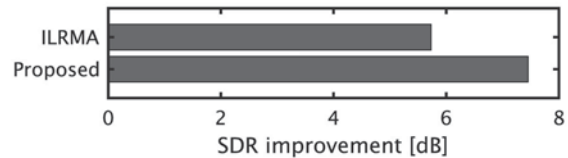


Fig. 1: Source-to-distortion ratio (SDR) improvement for each method.

1-11-3

1-11-3

木構造の階層的スパース性に基づく独立ベクトル分析

Independent vector analysis based on hierarchical sparsity of tree structure

☆岸本龍央(農工大)、松本和樹(早大)、山田宏樹、矢田部浩平(農工大)

背景 音には周波数帯域ごとに構造が異なるものがある
例: 音声の調波は低域に、歯擦音は高域に集中

提案 IVAに木構造を導入(周波数を階層的にグループ分け)

- ➡ 広い帯域にまたがる構造と
- 狭い帯域のみに現れる構造の両方を考慮

結果 木構造の導入が性能向上に寄与することを確認

- ➡ ILRMAと同程度の分離性能を達成

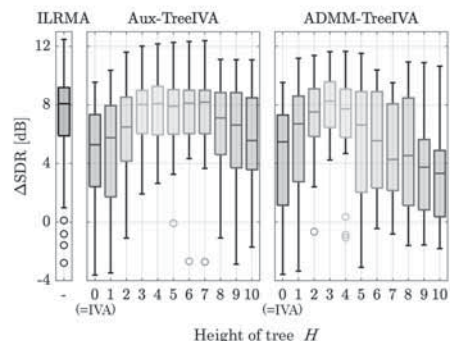


Fig. 1 ΔSDR of conventional ILRMA and proposed TreeIVA after 1000 iterations

1-11-4

1-11-4

Projection Backを距離射影で扱う 制約付き優決定ブライント音源分離

Constrained determined BSS realizing
projection back through metric projection

© 松本和樹(早大), 山田宏樹, 矢田部浩平(農工大)

従来法: 三段階の分離行列推定 (Fig. 1, 左)

白色化 / 分離行列の最適化 / Projection Back (PB)

- ▶ 最適化段階の分離音の持つ周波数特性の歪みが音源をモデル化する際に問題となり得る

提案法: PB後の分離音を直接扱う枠組み (Fig. 1, 右)

PBに関する凸制約付き最適化問題を近接分離法で解く

- 最適化の段階でPB後の分離音を扱う
- 制約集合 \mathcal{W}_{PB} への距離射影を利用することが近接分離法での安定した最適化に寄与する

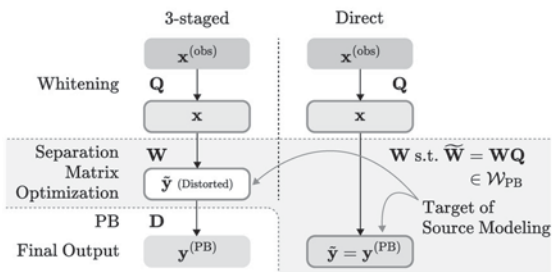


Fig. 1 3-staged (conventional) and direct (proposed) procedures

1-11-6

基底自動分配のための正則化を用いた 非負値テンソル因子分解による スポットフォーミング

Automatic basis allocation using attractor-based regularization
for NTF-based spotforming

© 綾野翔馬(香川高専), 李莉, 関翔梧(サイバーエージェント),
北村大地(香川高専)

◆スポットフォーミング

- ▶ 特定の領域にある目的音源を強調する技術

◆提案手法

- ▶ 複数ビームフォーマの出力の共通成分を抽出することで実現
- ▶ Fig.1に示すような非負値テンソル因子分解を適用
- ▶ 分配行列に対して正則化, 自動的なクラスタリングを導入

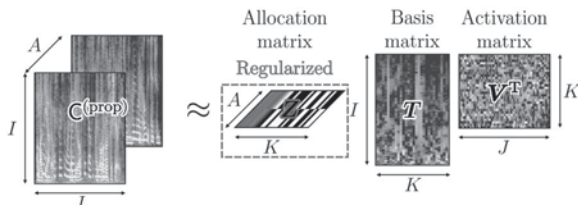


Fig.1: Model of nonnegative tensor factorization with regularization.

◆評価実験

- ▶ ハイパーパラメータの調整が容易であり, 基底数が十分であれば高い性能が得られる

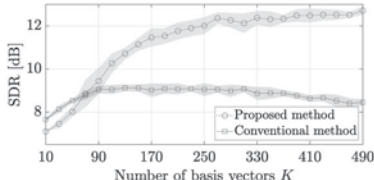


Fig.2: Plots of SDRs in simulation when $T_{\infty} = 0$ ms.

1-11-5

1-11-5 エイリアス成分のブライント音源分離による帯域復元の マイクロホン数増加による拡張

Extension of unaliasing based on blind source separation
by increasing the number of microphones

©中村優希, 中嶋大志, 小野順貴(都立大), 宮崎亮一(徳山高専)

◆我々は先行研究で, エイリアシングが生じた観測信号に対してブライント音源分離を適用し, 低域に混入したエイリアス成分を分離し高域を復元するアプローチを提案した.

◆本稿では, 連続時間信号に含まれる最大周波数成分に対し, サンプリング周波数が十分でなく, エイリアシングによって信号が2回以上折り返される場合において, 折り返し回数に応じて観測チャンネル数を増やすことで, 同じサンプリング周波数でもより高周波数帯域の成分を復元できることを示す.

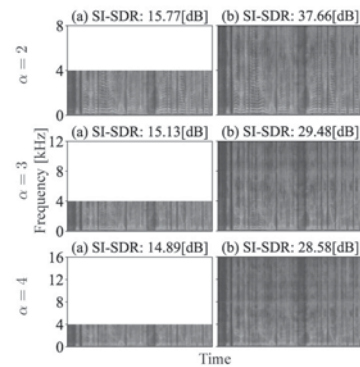


Fig.1 Examples of speech spectrograms.

(a) aliased observed signal, (b) restored signal by proposed method.

1-11-7

1-11-7 正負ラベルなし学習を用いた半教師付き 深層学習に基づくモノラル音声強調

Monaural speech enhancement based on semi-supervised deep learning
using positive-negative-unlabeled machine learning

☆小川諒, △米倉悠記, 伊藤信貴, 高宗典玄, 山岡洗瑛, 齋藤佑樹,
猿渡洋(東大)

◆従来手法: 教師付き深層学習に基づくモノラル音声強調

- ▶ 大量の雑音あり及び雑音なし音声の組(教師付きデータ)が必要
- 雑音なし音声の収録は静かな環境を必要とし高コスト

◆研究の動機: 教師付きデータが少量しかなくても, 低コストで大量に得られる雑音あり音声(教師なしデータ)を活用

◆提案手法: 半教師付き深層学習に基づくモノラル音声強調

- ▶ 正(positive: P), 負(negative: N), ラベルなし(unlabeled: U)データからの半教師付き学習法である正負ラベルなし(PNU)学習を利用
- ▶ 教師付きデータからP,Nデータ, 教師なしデータからUデータを取得し, PNU学習により各時間周波数成分をP,Nに分類

◆実験結果

- ▶ 提案手法(Proposed)は教師なし訓練データのサイズ(クリップ数) Q の増加に伴い音声強調性能 SI-SNRi*が改善し, $Q = 200, 400$ の場合に従来手法(Conventional)の SI-SNRi を 1.86, 3.03 dB 上回った

*Scale-invariant signal-to-noise ratio improvement

Table 1: SI-SNRi on the test data. Q is the number of noisy speech clips in the unlabeled training data

Method	Q	SI-SNRi (dB)
Conventional (PN)	0	3.52 (± 3.04)
Proposed (PNU)	1	1.32 (± 0.28)
Proposed (PNU)	200	5.38 (± 2.13)
Proposed (PNU)	400	6.55 (± 1.35)

1-11-8

1-11-8 DP マッチングと相互相関を用いた欠損を含む信号からの同期信号の生成

Synthesis of a synchronized signal from partially missing signal using dynamic programming matching and cross correlation

☆竹内駿人, 河村隆生, 小野順貴(都立大), 荒木章子(NTT)

◆目的

- 機器の故障や通信時のパケットロスなどにより、録音信号には欠損が生じる可能性がある。部分的な欠損を含む非同期録音信号からブラインドに同期信号を生成する手法を提案する。

◆概要

- 各録音信号を短時間フレーム分割し、フレーム間の相互相関関数のピークをコストとして、DP マッチングを適用することでフレーム間の対応関係を推定する。求めた対応関係に基づき、フレームのオーバーラップ加算をすることで同期信号を生成する。

◆実験

- 評価実験において、同期信号を欠損・伸縮が生じる前の信号との SI-SDR により評価し、提案手法の有用性を確認した。

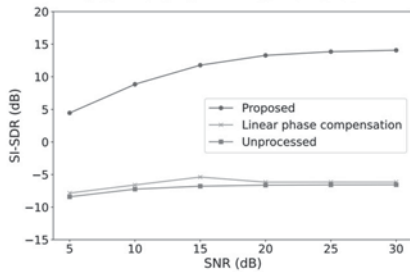


Fig. 1: Average SI-SDR for each SNR

1-11-10

1-11-10 人間とアバターとの対話システムにおける拡散性雑音下リアルタイム推定雑音を用いた Lombard 効果模擬音声合成のための検討

Lombard-Effect-Mimicking Processing for Speech Synthesis Using Real-Time Diffuse Noise Estimation in Human-Avatar Dialogue Systems

◎石川悠人, 武伯寒(東大), 中村友彦(産総研), 高宗典玄, 齋藤佑樹(東大), 高道慎之介(東大/慶大), 猿渡洋(東大)

◆人間を模したアバター/ロボットの対話システムに着目

- 雑音下ではアバター/ロボットの出力音声聞き取りにくくなり、またアバターとロボットの切り替わりで出力音声が不自然になる

◆雑音下で人間(ユーザ)とアバター/ロボットが対話する状況において、ユーザが聞き取りやすい音声を生成するフレームワークを提案

- 我々が以前提案した音声強調手法を活用し、ユーザの背景雑音をリアルタイムに推定
- 推定雑音に対して一定かつ十分大きなパワーとなるように出力音声をリアルタイムに調整
- さらにLombard効果の発話特徴量変化を出力音声上で部分的に再現し、より聞き取りやすい音声を生成

◆主観評価により、一定のゲインをかけたような従来手法と比べて雑音環境下でより自然な音声生成できることを確認

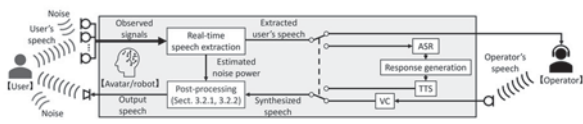


Fig. 1: Schematic of proposed framework for dialogue system

1-11-9

1-11-9 凸最適化に基づく構造付きスパース時間周波数解析のADMMによる高速化

Acceleration of convex-optimization-based structured sparse time-frequency analysis by ADMM

◎新井慶大(農工大), 松本和樹(早大), 山田宏樹, 矢田部浩平(農工大)

背景 離散ガボール変換による時間周波数(T-F)表現
広がり後段の処理に影響 ▶ スパース表現が望ましい

前回 構造付きスパース T-F 表現を得る凸最適化手法を提案

課題 主双対分離法(PDS)による最適化
▶ 収束までに多くの反復が必要

提案 交互方向乗数法(ADMM)による最適化
+ ネステロフの加速・適応ステップサイズの適用
▶ アルゴリズムの更なる高速化を実現

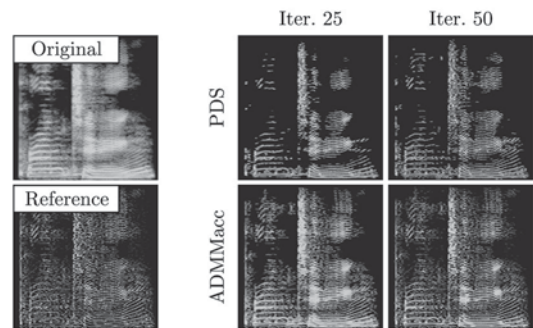


Fig. 1 Obtained T-F representations by PDS and accelerated ADMM

1-11-11

1-11-11 フレキシブルマイクを用いた騒音環境での音声入力

Flexible Diaphragm Microphone for Speech Input in Noisy Environment

◎平城裕隆(東京大院・学際), 三浦貴大, 金澤周介, 吉田学, 持丸正明(産総研), 厩本純一(東京大/SonyCSL)

◆導電布を用いたフレキシブルなダイヤフラムを用いたエレクトレットコンデンサマイクを制作し、騒音環境でも装着者の声のみを取得できるデバイス提案する。

◆マウスシミュレーターを用いて発声を模倣し、インパルス応答を計測することで評価した。

◆平面型のマイクでは正面や真後ろからの音が入力されやすいが、マスク型に湾曲することで装着者の声外部雑音よりも入りやすいことを周波数特性で明らかにした。

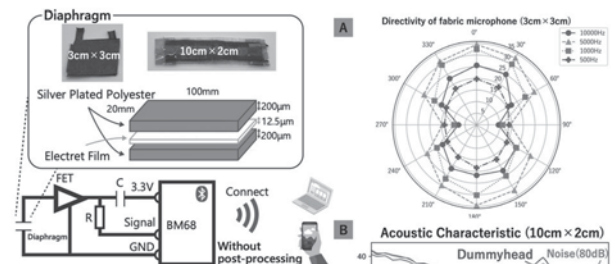


Fig. 1 (above) : Design of conductive fabric diaphragm for speech input.

Fig. 2 (right) : (A): Directivity of plate-type fabric microphone (B): Comparison of impulse response between curved microphone and noise

1-11-12

1-11-12 複数の障害物で構成された音響凸レンズを用いた距離選択集音

Selective distance sound collection utilizing acoustic convex lens composed of multiple obstacles

☆毛塚文芽, 高橋義典(工学院大)

- ◆距離選択集音とは、指向性に加え、特定の距離に位置する音源からの音を強調して集音する技術である。
- ◆本研究では、複数の球体障害物を領域毎に分布密度を中心から外側に減少させて充填した音響凸レンズを設計し、距離選択集音を検討した。
- ◆時間領域差分(FDTD)シミュレーションを実施し(Fig.1)、音響レンズのある場合の音圧レベルについて、レンズがない場合の音圧レベルからの差を計算したところ、4~8kHzの周波数帯域においては音響レンズの効果によって集音している様子がみられた(Fig.2)。

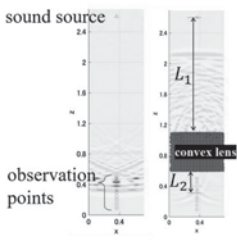


Fig.1 Simulated wavefront.

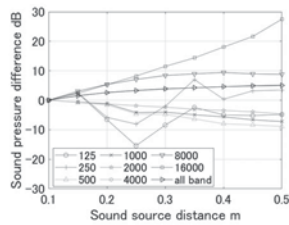


Fig.2 Sound pressure difference from without acoustic lens based on 0.1 m distance from the lens.

1-11-14

講演取消

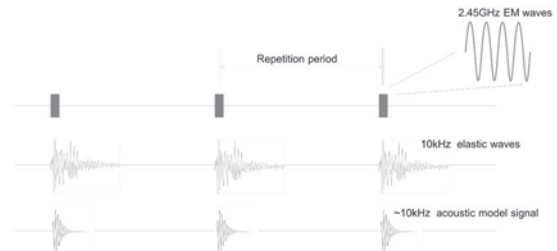
1-11-13

1-11-13 高周波パルス電磁界ばく露による聴覚的効果に関する実験的検証システムの開発

Development of an experimental system for the auditory effects of pulsed electromagnetic field exposure

○大久保寛, ハディ平和(都立大), △寺尾 安生(杏林大), 多氣 昌生, 畑 純一(都立大), 上田 麻里(神奈川工科大), 鈴木 敬久(都立大)

- ◆本研究では、パルス性電磁界ばく露が聴覚的に知覚される効果に関して、開発している実験的検証システムの報告を行う。
- ◆マイクロ波の照射実験のための電磁界パルスばく露システムの開発と併せて、繰り返しパルスによって知覚されると推定される音を音響シミュレーションによって作成し、音響信号による聴覚的な知覚特性を並行して行うことにより、閾値特性についてその関連性を明らかにする。
- ◆下図に、マイクロ波聴覚効果におけるマイクロ波照射と繰り返しパルス音の発生機構の概略を示す。上段は、ばく露されるマイクロ波を示す。中段は、熱弾性波シミュレーションによって得られた頭部内での音圧信号波形である。下段はモデル信号波形である。



1-Q-1

1-Q-1 かるまい文化交流センター (宇漢米館)の音響設計

Acoustical Design and Characteristics of Karumai Cultural Exchange Center

◎荒田耕平, 北村浩一(NHK-TECH)

- * 広い残響可変幅 (空席時・500Hz/約1.8~1.0秒)
- * 赤松無節仕上の側方拡散壁
- * 木リブ仕上げの天井反射板と軸回転片持ち式の側方反射板

Table.1 ACOUSTICAL ATTRIBUTES

Attribute	1/1 OCTAVE BAND CENTER FREQUENCIES IN HERTZ					
unoccupied	125	250	500	1000	2000	4000
RT [sec]	1.81	1.79	1.77	1.64	1.47	1.38
EDT [sec]	1.66	1.74	1.63	1.61	1.45	1.36
C ₈₀ [dB]	-3.5	-1.0	-0.7	-0.7	-0.1	0.0
Ts [msec]	153	135	120	120	112	110

with stage reflectors set for concerts

Fig.1 HALF PLANS

Fig.2 LONGITUDINAL SECTION

Fig.3 REVERBERATION TIME

Fig.4 WALL DIFFUSERS

Seating Capacity 400
Room Volume 4,269 m³
Surface Area 1,742 m²

REVERBERATION TIME IN SEC.

with stage curtains lowered
with stage reflectors set for concerts
unoccupied

63 125 250 500 1K 2K 4K 8K
1/3 OCTAVE BAND CENTER FREQUENCIES IN HZ

【MAIN FINISH MATERIALS】
A : Fiber Reinforced Gypsum Boards 6t+Decorative Boards 6t
B : Gypsum Boards 9.5t+Pine 50φ×15t
C : Gypsum Boards 9.5t+Plywood 6t
D : Wood Rib 18×20 or 35 or 45+Glass Cloth+Glass Wool 50t
E : Wood Rib 60×120 or 180
F : Gypsum Boards 9.5t+2t+Decorative Boards 6t
G : Glass Cloth+Glass Wool 100t

1-Q-2

1-Q-2

水戸市民会館の音響設計

Acoustical Design and Characteristics of Mito City Civic Center

○鈴木航輔, 福地智子(永田音響設計)

◆水戸市民会館は茨城県水戸市に建設された公共施設で、2023年7月に開館した。大ホール(2000席)、中ホール(482席)、小ホール(平土間)、大小さまざまな会議室などで構成され、コンサートやコンベンションで利用されている。本稿では大ホールと中ホールの室内音響計画について報告する。

◆大ホール

- ▶ フライタワーを備えた多目的ホール。客席は1~3階席の3層構成。
- ▶ ポップス等の拡声設備を用いたツアーコンサートを招致できる設備と収容人数を持ち、オーケストラや吹奏楽など生音のみの公演に利用される場として計画された。
- ▶ 壁面がコンクリートである。フラッターエコーを防止し聴感的に鋭い反射音を和らげるため、小叩き仕上げの上に梅の花を模したGRCの曲面パネルを取り付けることとした。

◆中ホール

- ▶ 長方形の多目的ホール。
- ▶ 生音のコンサートや拡声設備を利用した演劇など、市民が利用しやすいホールとして計画された。
- ▶ 壁面はフラッターエコーを防止するため石膏ボードと木練付仕上げによる矢羽根模様の凹凸を設けた。
- ▶ 同規模のホールが隣接する水戸芸術館に2つあり、残響時間はこれらのホールよりやや短めである。

1-Q-4

1-Q-4 四万十市総合文化センター「しまんとぴあ」の音響設計

Acoustical design and characteristics of Shimantopia

○和田竜一, 小野朗(永田音響設計)

2024年4月、高知県の四万十市に開館した「四万十市総合文化センターしまんとぴあ」は、大小2つのホールと各種スタジオからなる複合文化施設である。

■ 遮音計画 : 各室は交流ロビーと呼ばれる共用空間を中心として、廊下や緩衝スペースを介して配置した。そのうえで、小ホール周囲の構造的エキスパンション、スタジオやリハーサル室に防振遮音構造を採用して、出来るだけ同時使用が可能となるようにした。

■ 大ホール(805席)の室内音響計画 : 大ホール形式は、メインフロア、正面バルコニー席と客席側方部分のサイドバルコニー席からなるプロセニウム型ホールである。室形状は六角形平面で、舞台と客席を近づけるため、やや幅広(最大30m)である。視覚/音響の両面から舞台への親密感を得るため以下のような検討を行った。

- ・舞台/客席の天井高をなるべく高く確保し、天井反射板や客席天井を下向きの凸曲面とした。
- ・側壁及び後壁を有効な反射面として活用できるよう、その形状、角度を検討した。
- ・サイドバルコニー下部壁面を内傾させ初期反射音を客席中央に到達させた。



大ホール断面図

1-Q-3

1-Q-3 KAIT 広場の音響測定とホールマネジメントエンジニア(HME)育成への応用*

Acoustic measurements at KAIT Plaza and their application to Hall Management Engineer (HME) training

☆大石まなか(神奈川工科大), 田中邦和(サクソ奏者), 西口磯春, △高尾秀伸, △板子一隆, 上田麻理(神奈川工科大)

- ◆神奈川工科大学におけるKAIT広場は、2020年に竣工された広さ約4100m²の無柱の半屋外空間の多目的施設である。
- ◆その開放的な空間設計が特徴で、特に、天井の58カ所の開口部は、音響特性に大きな影響を与えることが予想される。
- ◆本研究では、KAIT広場の音響特性を明らかにするために、残響時間の計測や遮音性能の評価を行った。また、KAIT広場の音楽的な利活用の価値を高めるため、プロの演奏家を対象に演奏しやすさをアンケートにより評価してもらった。



Fig. 対象空間としたKAIT広場

1-Q-5

1-Q-5 残響時間の弁別閾と音響的素養の関係性についての基礎的検討

Fundamental study of relationship between the discrimination threshold of reverberation time and acoustical interest and experience

◎原田和典(岡山県立大)

- ◆本検討では極限法を用いて、残響時間の弁別閾値を求め、その結果と、被験者の楽器経験や音へのこだわり(以下、音響的素養と呼称)の関係について検討した。また残響時間の違いを明確に認識できる閾値についても検討した。
- ◆ヴァイオリンの演奏音と女性の朗読音声、インパルス応答を試験音とし、フィルター加工を施したインパルス応答を畳み込むことで0.3~2.4sの残響を持つ0.1s刻みの刺激を用いた。
- ◆残響時間の弁別閾値について、既往研究の値のおよそ2倍の値でその違いをはっきり異なると判定された。
- ◆楽器経験や音楽への没頭経験がある場合被験者はより細かく違いを判断していた。

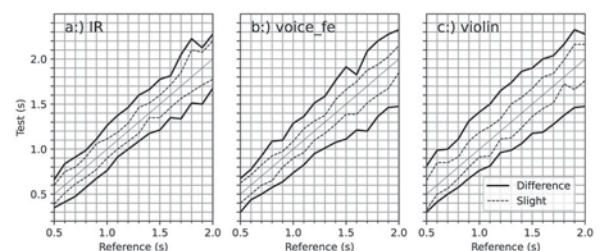


Fig.1: Average Thresholds by Sound Source

1-Q-6

1-Q-6 劇場の天井耐震改修事例
—改修工法に応じた音場の最適化—
Theater ceiling earthquake resistance renovation example
- Optimizing the sound field according to the renovation method -

○日根野翔太, 宮崎秀生(ヤマハ)

- ◆2024年現在でも全国各地で劇場の天井耐震改修が行われている。
◆劇場における天井耐震改修工事の音響上の目的は、「音場保存」と「音場改善」に分けられるが、改修後の音場を最適化するという点においてはどちらも共通している。
◆特定天井改修には複数工法が存在し、それぞれ音響的な特徴が異なるため、各工法によって生じる音場への影響などを把握しておくことが音場の最適化を実現する上で重要である。
◆本稿では、各天井改修手法の特徴を述べるとともに、各工法で実施した音響検討内容について紹介する。

表1 天井耐震改修の工法種別と特徴

Table with 4 columns: 改修種別, 改修概要, 工法, 特徴. It lists various renovation methods like '既存構造の補修', '天井構造の最適化', etc.

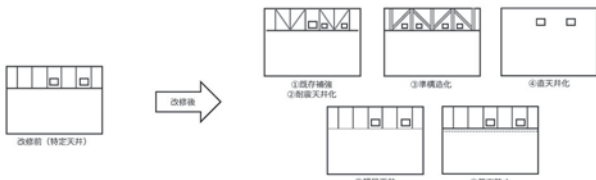


図1 天井耐震改修概要図

1-Q-8

1-Q-8 マイクロスリットを用いた音響メタサーフェス吸音体の簡易設計法
-共鳴器群の吸音率の設計に関する検討-
Simple design method of acoustic metasurface absorbers using microslits
-Absorption design of component resonators-

☆木村祐希, 奥園健(神戸大院・工)

- ◆準完全吸音を実現するマイクロスリットを用いた音響メタサーフェス吸音体(MeAMS吸音体)は、共鳴器群を完全吸音からかけ離れた低い吸音率ピークをもつ不完全共鳴器として構成することで、目標の吸音特性を頑健に設計でき、薄型化可能である。
◆上記に限られた共鳴器数の場合で明らかにされているため、本研究では、共鳴器数を複数設定して、共鳴器数によらず不完全共鳴器とする理論設計が、目標とする吸音特性の実現し易さや構造の薄型化に効率的であるかを数値解析、並びに、音響管法を用いて検証した。
◆さらに、共鳴器数の変化により共鳴器群が全体の吸音にどのように寄与しているかを数値的に明らかにした。



Fig.1: Samples of acoustic metasurface absorbers using microslits.

1-Q-7

1-Q-7 音楽練習室の音響設計に関する研究
—三味線練習室を対象としたケーススタディー—

Study on acoustic design of musical practice room: A case study for shamisen practice room

○岡本則子, 富来礼次, 大鶴 徹(大分大), 渡部奈津実(大分大院)

- ◆既報に引き続き、三味線練習室を有する木造住宅の建替え計画の過程を示し、建替え後の練習室の音響性能の測定結果を報告する。
◆本研究では、これまで研究事例の多くない津軽三味線を対象とするため、まず、津軽三味線の演奏を収録し、周波数特性を確認した。
◆設計目標値が明確でないため、奏者がこれまで演奏に問題を抱えることのなかった旧住宅の練習室(Room A)を参考に計画することとし、Room Aのインパルス応答の測定および内装材の吸音特性のin-situ測定を行った。また、参考として、奏者が不快と感じた室(Room B, C)においてもインパルス応答の測定を行った。
◆確認のため、建替え後の室(Room D)で実測と三味線の演奏を行った。Fig. 1に示すとおり、Room Dの残響時間および平均吸音率は目標値と概ね一致した。犬皮の津軽三味線使用時では問題がないことが聴感的にも確認された。

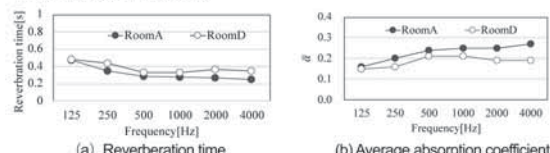


Fig.1: Comparison of reverberation time and average absorption coefficient between Room A and Room D.

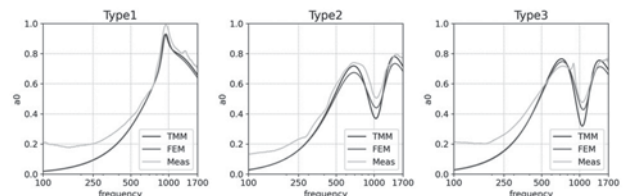
1-Q-9

1-Q-9 マイクロスリット共鳴器を埋め込んだメタポーラス吸音体に関する研究
—共鳴器の埋め込み位置を考慮した理論モデルの開発—

Study on metaporous sound absorbers with embedded microslit resonators
-Development of theoretical models considering resonators position-

☆藤井倫太郎, 奥園健(神戸大院・工)

- 共鳴器配置の異なるマイクロスリット共鳴器を多孔質材に埋め込んだMeMP吸音体(Type1~Type3)に対して理論モデルを提案する。
■理論モデルは凸形状の分岐管路構造をもつ多孔質層とマイクロスリット共鳴器で構成される単位構造を伝達マトリクス法(TMM)によって記述し、熱粘性音響理論を用いたFEM解析、3Dプリンタによるプロトタイプを用いた音響管実験との比較によってその妥当性を検証した。
■FEM解析では、音響インテンシティ分布により分岐管路構造の音響伝搬のモデル化の妥当性を確認し、また多孔質材の流れ抵抗率の変化に対しても良好な一致を示した。
■全てのTypeでFEM、実測(Meas)に対して妥当な近似が得られ(下図)、提案する理論モデルの妥当性が示された。



1-Q-10

1-Q-10 反射音の重畳を利用した コミュニケーション補助壁の提案

Communication assist wall utilizing superposed reflected sound

☆今田彩文(工学院大), 高橋義典(工学院大)

- ◆ 反射音で音声の伝達を補強する壁面設置型のデバイスを提案する。
- ◆ 複数の反射板で構成したコミュニケーション補助壁と音源・受音点を焦点とした楕円の補助壁について FDTD シミュレーションを実施し、どちらも反射波が受音点に届いていることが確認できた (Fig. 1-2)。
- ◆ コミュニケーション補助壁を試作し、無響室で通常の壁面との比較実験を行った(Fig. 3)。
- ◆ 実験の結果、直接音の到達から 1 ~ 2 ms 以下の間に複数の反射音が重畳し、通常の壁面よりも 2.5 dB の音圧の上昇が確認された (Fig. 4)。

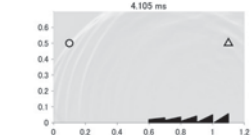


Fig. 1 Reflected waves from a communication assist wall composed of multiple reflectors. O: observation point, Δ: sound source.

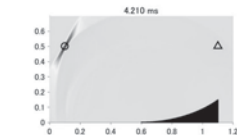


Fig. 2 Reflected waves from an elliptical communication assist wall. O: observation point, Δ: sound source.

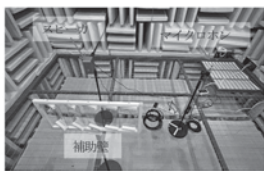


Fig. 3 Experimental arrange in an anechoic room

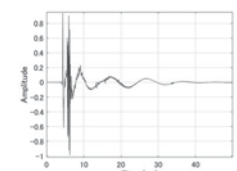


Fig. 4 Each measured impulse response (Red : plate , Blue : communication assist wall composed of multiple reflectors)

1-Q-12

1-Q-12 編物の吸音特性に関する基礎的研究 —実験的検討—

Sound absorption characteristics of knitted fabrics -An experimental study-

☆高橋響, 阪上公博(神戸大院・工)

- ◆ 編物の吸音特性に関する研究は特定の用途に留まるものが多く、先行研究が乏しいため、建築内装材としての利用を目的とした基礎的な実験的検討を行なった。
- ◆ 作成した編物試料の垂直入射吸音率、比音響インピーダンス比を測定し、吸音特性を検討した。Fig. 1の通り、全体を通して多孔質型吸音の傾向が見られ、流れ抵抗、部材厚、面密度などのパラメータが吸音特性に寄与していることが分かった。
- ◆ 既存の予測理論 (Miki, Allard, Kato, Allard+Miki) と比較した。Fig. 2の通り、定量的に一致を示すものはなかったが、多孔質型吸音の予測理論である Miki モデル、Kato モデルでは定性的に近い傾向が見られた。Allard モデルや、Allard モデルの円筒管以外の試料部分に Miki モデルを当てはめた Allard+Miki モデルは、共鳴型吸音を含んだ予測理論だが、編み形による孔は縦横に囲まれているため、共鳴による寄与は少なく、実測値からはやや外れた結果となったと考えられる。

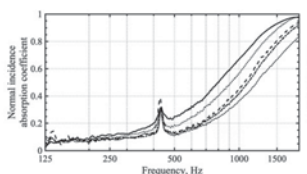


Fig. 1: Measurement result (wool(7), back cavity 20mm)

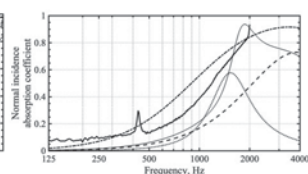


Fig. 2: Prediction result (acrylic(7), stockinette stitch, back cavity 20mm)

1-Q-11

1-Q-11 複数の障害物で構成された 音響ディフューザの拡散性の評価

Evaluation of diffusion effect of acoustic diffuser composed of multiple obstacles

☆増川英寿, 高橋義典(工学院大・情報)

- ◆ 寸法と周波数の関係について明らかにするために複数の円盤で構成された音響ディフューザを設計した(Fig. 1)。
- ◆ FDTD シミュレーションの結果から、正規化拡散係数を計算した(Fig. 2)。2 kHz以上で今回提案した音響ディフューザの方が従来型より正規化拡散係数が大きかった。
- ◆ 168 m³の矩形残響室に拡散体を敷き詰め時の残響時間を計算したところ、提案型は従来型より短かった(Table 1)。障害物の個数を変えることで吸音力を制御できると考えられる。

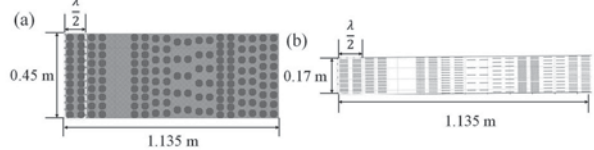


Fig. 1 Proposal type of acoustic diffuser (a) front view, (b) top view

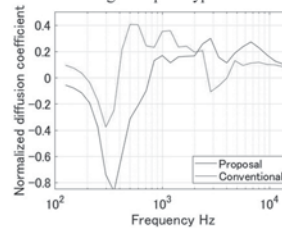


Fig. 2 Normalized diffusion coefficient

	Reverberation time(s)
Proposal	0.781
Conventional	9.26
Rectangular reverberation room	13.49

Table 1 Reverberation time

1-Q-13

1-Q-13

1-Q-13 不均一な背後空気層を持つ 板振動型吸音体に関する基礎的研究

A basic study on panel sound absorber with an air back cavity of non-uniform depth

☆千田純理, 阪上公博(神戸大院・工)

- ◆ 本研究では、弾性板と背後壁が平行でなく、厚さが不均一な背後空気層を持つ板振動型吸音体の吸音特性を、COMSOL Multiphysics を使用した数値解析によって検討した。2次元モデルにおいて、背後壁に傾斜角を与えて傾斜させて、容積一定で空気層厚を不均一とした。
- ◆ 背後壁の傾斜角を変化させても、ピーク周波数やピーク吸音率に変化がほとんど見られなかったため、不均一な背後空気層の板振動型吸音体の吸音特性は背後壁の傾斜角に依存しないと考えられる。
- ◆ 実際に不均一な背後空気層を持つ板振動型吸音体を作成し、音響管により垂直入射吸音特性を測定し、数値解析の結果と比較した(Fig.1)。解析における弾性板の端部条件は、自由端と固定端である。この結果から、本論文の数値解析による知見は、実測結果と定量的には一致していないが、定性的な傾向には一致が見られることが分かった。

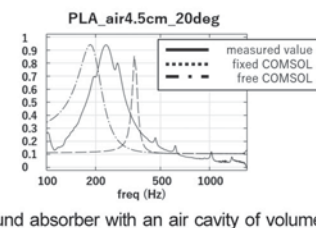


Fig. 1: Panel sound absorber with an air cavity of volume 4.5 cm x 10 cm. Normal incident sound absorption coefficient at an inclination angle of back wall $\theta=20^\circ$. Straight line: measured results; dotted line: numerical results assuming fixed ends; dashed line: numerical results assuming free ends.

1-Q-14

1-Q-14 数値シミュレーションによる異形孔を有する微細穿孔板の吸音特性に関する検討

Study on sound absorption properties of micro perforated panels with irregularly shaped holes by numerical simulation

☆阿部桜子, 阪上公博(神戸大院・工)

- ◆背景

近年、微細穿孔板(MPP)の意匠性が注目され、異形孔を有するMPPに代表される美的に優れたデザインの商用製品が増加している。また、低コストで製造されたMPPの穿孔は、非円形で不規則な形状となる場合がある。
- ◆概要

本研究では、異形孔(正方形孔、正三角形孔)を有するMPPと円形孔と異形孔が混在するMPPの吸音特性について、COMSOL Multiphysics® Ver. 6.2の圧力音響モジュールによる数値シミュレーション(モデルをFig.1に示す)を用いて、従来のMPPの吸音特性との比較による検討を行う。
- ◆結果
 - ・正方形孔MPPの場合は、孔の流れ抵抗率、孔の流れ抵抗率と開孔率、正三角形孔MPPの場合は、孔の流れ抵抗率と開孔率を円形孔MPPと一致させることで円形孔MPPと同等の吸音特性が得られた。
 - ・これらの特性量の一致を用いることで、円形孔と異形孔が混在するMPPによって、円形孔のみで構成されるMPPと同等の吸音特性が得られることが分かった。

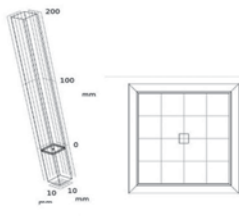


Fig.1: Simulation model

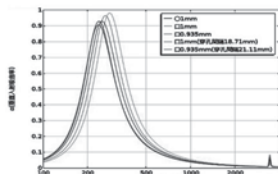


Fig.2: Sound absorption properties of each of the rectangular/triangular hole MPPs with each characteristic quantity matched to a circular hole MPP

1-Q-16

1-Q-16 矩形ソニックブラックホール吸音体に関する基礎的研究

Fundamental research on rectangular sonic black hole sound absorbers.

☆上辻誠五, 奥園健(神戸大院・工)

- ◆本研究では矩形断面を持つソニックブラックホール(SBH)吸音体を提案している。
- ◆矩形SBH吸音体3Dモデルの作成方法と3Dプリンタでの造形方法について提案する(図1)。
- ◆数値解析並びに音響管実験を用いて、円形SBHと矩形SBHの吸音特性を比較し、提案する矩形SBHは円形SBHとほぼ同等の吸音特性を得られることを明らかにする(図2)。

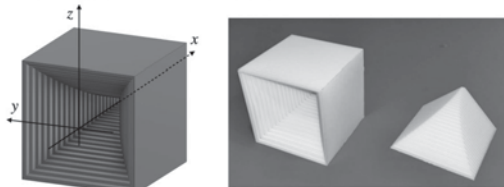


図1: 矩形SBHの3Dモデルと、3Dプリンタで作成した試験体と構成ユニット

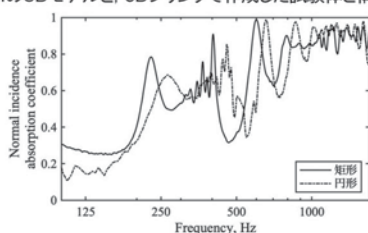


図2: 矩形・円形SBH吸音体の垂直入射吸音率(実験)の比較

1-Q-15

1-Q-15 矩形2次元音場における深層学習を用いた吸音率推定の基礎的検討

Data-driven estimation of sound absorption coefficients in the rectangular 2D sound field

☆大河佑里子, 新井大斗, 佐藤元, 池田雄介(東京電機大)

- ◆背景
 - 近年、測定と数値シミュレーションを併用した、深層学習ベースの吸音率推定手法が提案されてきた
 - 先行研究では、室内全体の平均吸音率の推定や個々の壁面の吸音率の推定が行われている
- ◆提案手法
 - 少数のマイクロホンの音圧時間信号から、MLPを用いて矩形2次元音場における4壁面それぞれの吸音率を推定
- ◆実験
 - 中央値では十分な精度で吸音率の推定可能
 - 分散が大きく、推定精度の安定性に課題

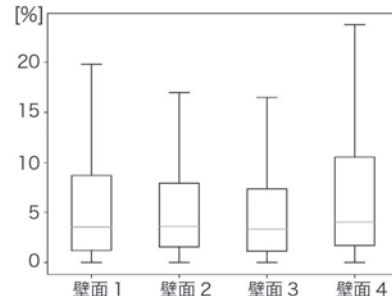


Fig.1: Relative errors of sound absorption coefficients

1-Q-17

1-Q-17 物理モデルと深層学習を用いた壁面の音響インピーダンス推定

Acoustic impedance estimation of wall surfaces based on deep learning using physical models

☆新井大斗, 佐藤元, 池田雄介(東京電機大)

- ◆背景
 - 実環境に即した音場シミュレーションを行なうためには、境界条件の把握が必要
 - 近年、計測の効率化を目的とした深層学習を用いた吸音率の推定が提案
- ◆提案手法
 - 損失関数に物理モデルを導入したニューラルネットワークを用いて測定した伝達関数から音響インピーダンスを推定
- ◆実験
 - 相対誤差の中央値において20%程度で吸音率の推定が可能
 - 物理モデルの有効性の検証が更なる課題

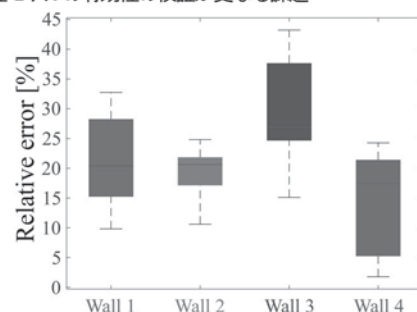


Fig.1: Relative error of estimated acoustic impedances

1-Q-18

1-Q-18 物理法則に基づいた深層学習による 正方形室内音場の推定

Sound field estimation in square rooms using deep learning based on physics information

☆佐藤元, 池田雄介(東京電機大)

◆背景

- 近年, より高速に音場をシミュレーションするために, 深層学習を応用した音場推定手法が提案されている。
- 一方で, 多くの手法において任意点の音圧を求めるためには, 別途データ補間を行う必要があった。

◆提案手法

- 損失に Helmholtz 方程式を用いることで, 物理法則に基づいて, モデルを学習。
- PINN の活性化関数に正弦関数を用いた深層学習モデルで, 空間座標と部屋形状の情報から室内音場の複素音圧を推定。

◆結果

- データのみで学習したモデルと比較して SNR で最大で約 4.9 dB 推定精度を改善。

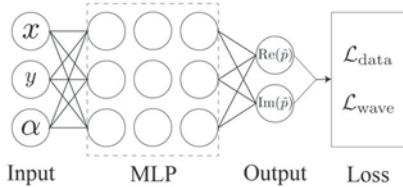


Fig.1: Architecture of the proposed method.

(x, y: coordinate, α : room shape information, \hat{p} : estimation sound pressure)

1-Q-20

1-Q-20 プレナムドアの遮音性能向上に関する 数値解析的検討

—吸音処理とMPPの挿入による効果—

Numerical Analysis on Improvement of Sound Insulation Performance of Plenum Doors: Effect of sound absorption and MPPs

☆兒玉和紗, 阪上公博(神戸大院・工)

- ◆プレナムドアの遮音性能を, 吸音処理を行う場合について, 有限要素法を用いた数値解析によって検討した。また, プレナムドアに MPP を挿入することを考え, その効果的な挿入方法について検討を行った。
- ◆プレナムドア内部に吸音処理を行うことで, その遮音性能の向上が可能であることが示された。その効果は吸音率と面積に依存する。
- ◆MPP を内部に設置することで, 対策目標の周波数で音響透過損失を増加させることが可能であった。本研究では, MPP の面積や設置条件により MPP の効果に差があることが示された (Fig.1)。設置方法については十分な検討が必要である。

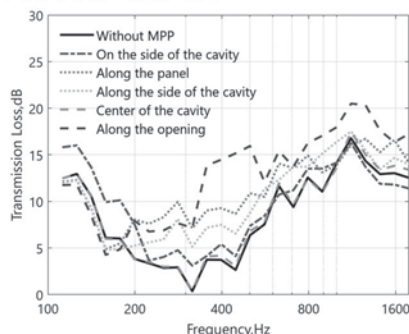


Fig.1: Transmission loss of a plenum door when an MPP is installed inside.

1-Q-19

1-Q-19 木製音響リブ拡散体の会議室適用例

Application example of wooden ribbed acoustic diffuser in conference room

○中津成博, 上田泰孝(安藤ハザマ), 俣野祐美, 大谷美友(大建工業)

- ◆内装木質化と室内音響調整を同時に実現する後付けパネルとして木製の交差リブ拡散体(以下,CRD)を作成し, 小規模会議室(3.13m×4.96m×2.5m(H))に実装した(Fig.1)。
- ◆小規模会議室内で CRD 設置前後のインパルス応答測定を行い, CRD の室内音響調整効果を検証したので報告する。
- ◆CRD 設置により 1000Hz 帯域以降で受音点位置によらず残響時間の低減がみられた(Fig.2)。
- ◆音声の明瞭度指標である D50 および STI についても残響時間同様に CRD 設置により向上することが確認された。

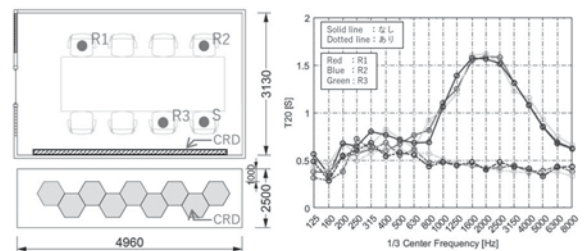


Fig.1: CRD placement and sound source/receiving point (Sound Source: S, Receiving Point: R1,R2,R3)

Fig.2: Reverberation Time(T20)

1-Q-21

1-Q-21 商店街環境におけるサウンドスケープの 最適化手法に関する研究

聴覚要素と視覚要素が相互に及ぼす影響

Soundscape in a shopping district environment

Research on optimization methods

Effects of auditory and visual elements on each other

○近藤真美, 許載永(奈良女子大学)

- ◆はじめに
本研究は商店街環境のサウンドスケープに焦点を当て、聴覚・視覚要素が訪れる人々の知覚や行動にどのような影響を及ぼすのかを明らかにすることを目的とする。さらに聴覚要素と視覚要素が相互に及ぼす影響についても検討を行う。評価対象地は日本 4ヶ所、韓国 4ヶ所で条件の提示は全て VR 上で行う。
- ◆結果・考察
印象評価、音源認知度評価ともに音源のみ提示(以下 AO)した場合と、音源と映像を同時に提示(以下 AV)した場合において結果に差が見られた。印象評価では評価地によって結果が異なり、傾向が見られなかったが、音源認知度評価においては AO よりも AV の方が騒音の認知度が低下する傾向が見られた。このことから聴覚情報に視覚情報が加わることで騒音による不快感が軽減されることが示唆される。

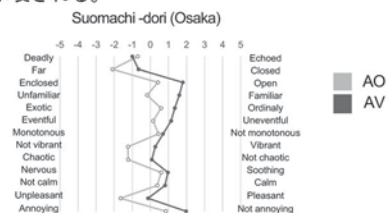


Fig1. Impression evaluation

1-Q-22

1-Q-22 バーチャルリアリティ技術を用いた音源の距離知覚の心理実験法に関する一検討

A study on psychological experiment methods for distance perception of sound sources using virtual reality technology
○佐藤逸人, △阿部瑞穂(神戸大院・工学研)

- ◆バーチャルリアリティ (VR) 技術を用いた心理実験について、視聴覚刺激の制御方法や実験参加者の回答方法の自由度が広がるという利点に着目し、音の距離感を回答させる心理実験法について検討した。
- ◆実験参加者のタスクは、VR空間において音源位置が視覚的にも提示される「基準刺激」と、音源位置が視覚的には提示されない「比較刺激」を自由に切り替えながら聞き比べて、比較刺激の音像位置に球を配置することとした。
- ◆共通の頭部伝達関数を用いたヘッドホン提示による実験では、音像方向の定位精度の低下や頭内定位の問題は避けられないが、音圧レベルと残響音が距離知覚に及ぼす影響について、過去の知見を再現する結果が得られた。また、室内の視覚情報が距離知覚に影響する可能性も確認することができた。

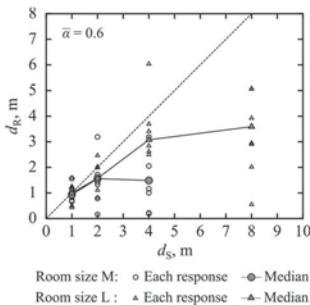


Fig. 1 Relationship between horizontal distance to sound source (d_s) and horizontal distance to responded location (d_r). Circles and triangles show the results for room size M (11 m × 13 m × 3 m) and room size L (22 m × 26 m × 3 m), respectively.

1-Q-24

1-Q-24 波動音響解析による室内音響予測の音響材料モデリングに関する研究—室容積の違いによるランダム入射吸音率モデルの適用性検証—

Acoustic material modeling for wave-based room-acoustics simulation-Applicability of random-incidence absorption coefficient model for different room volumes
☆田尻加奈子, 奥園健(神戸大院・工学研)

- ◆本研究では、波動音響解析による室内音響の予測において、音響材料モデリング手法として実用性の観点から多用される吸音率モデルの適用性をよりの確かな波動モデルとの比較から検証している。
- ◆本稿では、5種類の現実的な内装条件の普通教室 (Fig.1) を対象とした検証結果について報告する。
- ◆結果として、先行研究において小会議室を対象に明らかにした吸音率モデルの誤差傾向がより室容積の大きな普通教室でも成立することを確認した。

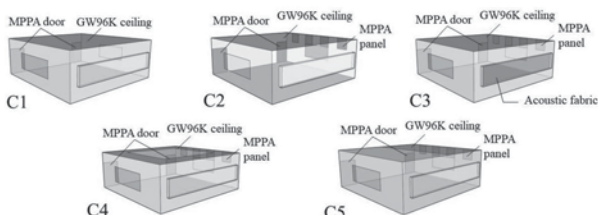


Fig 1 The classroom model @C1-C5

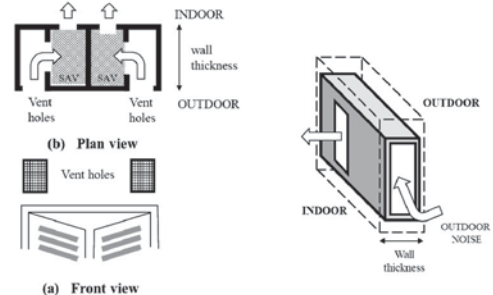
1-Q-23

1-Q-23 防音機能を有する通気口の音響特性の計算

Calculation of acoustic performance of Soundproofing Air Vents
○西村壮平(熊本高専), 西村勇也(熊本高専)

- ◆近年、東南アジアの発展途上国においては道路交通騒音をはじめとする騒音問題が年々深刻化されているにも関わらず、それに対する対策がほとんど行われていない。一方、東南アジアの住宅用通気口は建物の外壁に建物の大きさに応じた換気口が複数設けられている形状が一般的であるため、換気を行うと、同時に騒音もそこを通過して室内に伝搬してしまう。この問題を解決するために、本研究では換気と防音の両方の機能を兼ね備えた通気口を提案する。

防音通気口は防音と換気という二つの相反する性質を併せ持つ必要がある。そのためには、まず防音ユニット内の音波伝搬および騒音の要因となる平面波音圧成分と高次波音圧成分の発生メカニズムを明らかにする必要がある。本報告ではこれらの発生メカニズム、およびその防止方法を理論計算と実験検証の双方から明らかにする。



The front view and plan view of proposed SAV Soundproof Air Vent

1-Q-25

1-Q-25 時間領域有限要素解析による室内音場の粒子速度算出に関する研究—RC 矩形室を対象とした3軸音圧-粒子速度センサを用いた実測値との比較—

Study on particle velocity calculation by time domain finite element sound field analysis
- Comparison with measurement using three-dimensional pressure-particle velocity sensor in RC rectangular room -
☆境真佐樹(大分大院), 富来礼次, 岡本則子, 大鶴徹(大分大), 菅野将郎(大分大院)

- ◆時間領域有限要素解析 (TDFEM) により得られた音圧から粒子速度を近似し、境界要素法と比較し、よく一致することを示した。
- ◆RC造矩形室を対象に、TDFEMにより得られた粒子速度の初期応答と減衰曲線を3軸音圧-粒子速度センサによる実測値と比較した。
 - ✓ TDFEM と実測による粒子速度の初期応答は方向によらず、よく一致した。
 - ✓ TDFEM により得られた音圧の残響時間と実測値が近似するよう、境界に実インピーダンスを与えた際、粒子速度の減衰曲線は、方向によって実測値と近似しない場合があることが示された。

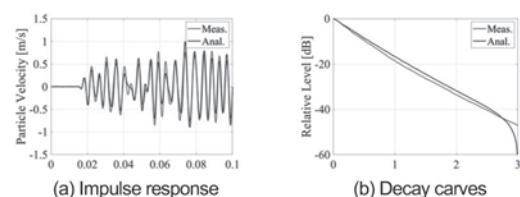


Fig.1 Comparison of x-dimensional particle velocity v_x between TDFE-analysis and measurements in 1/1oct. with f_c of 250Hz.

1-Q-26

1-Q-26 産業機器への適用を可能とした膜振動とヘルムホルツ共鳴を利用した広帯域吸音音響メタマテリアル

The evaluation of the developed sound absorber using the membrane vibration and Helmholtz resonator for implementation to industrial facility.

◎後藤 達彦, 江波戸 明彦, 西村 修 ((株) 東芝)

産業機器への適用を可能とする改良構造を示し、音響管評価、複数ユニット評価を実施した結果を示す。

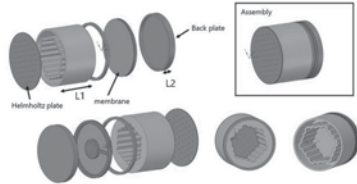


Fig.1: The structure of the proposed sound metamaterial.

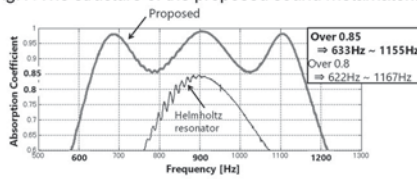


Fig.2: Measured absorption coefficient.

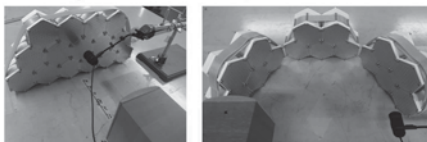


Fig.3: Evaluation with using multi-unit.

1-Q-28

1-Q-28 背後空気層を持つ有孔板を用いた家屋内低周波騒音の低減対策—模擬家屋を用いた実験的検討—

Measures to reduce low-frequency noise in the test house using sound-absorbing materials

◎土肥哲也, 岩永景一郎, 鈴木肇 (小林理研)

- ◆家屋内で簡便に実施可能な低周波騒音の低減対策方法を模索する目的で、家屋内に吸音体を後から設置する方法で家屋内の低周波騒音がどの程度低減するか模擬家屋(図1)を用いた実験的手法により検討した。
- ◆吸音体として背後空気層を持つ有孔板(高さ90cm, 横幅90cm, 奥行約30cm, 6面体の前面のみに直径6mmの孔: 図2)を12個製作した。市販の木板を使用し、1個の材料費は5000円と安価でDIY製作が可能な仕様とした。
- ◆吸音体を設置した場合の100Hzの床面SPL平均値は、無対策条件に比べて約10dB低減し、屋内における低周波騒音が低減する可能性が示唆された。



Fig.1: Test building

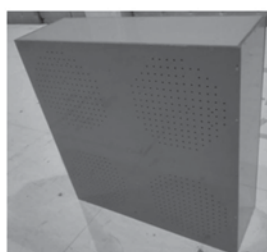


Fig.2: Sound absorber

1-Q-27

1-Q-27 折曲げ型簡易共鳴箱を用いた屋内低周波音の低減

Low-Frequency Noise Reduction in Residential Houses Using Folded Quarter-wave Resonators

◎長井健一郎, 久保凱, 生沼秀司, 榎本俊治, 石井達哉 (JAXA)

家屋内へ到達する低周波音の低減対策として、簡便に設置できる吸音デバイスの検討を行った。1/4波長の共鳴管をベースにした折曲げ型共鳴箱を考案し、汎用シミュレーションソフトウェア(COMSOL Multiphysics)で形状設計を行った。この吸音デバイスをダンボールで製作し、模擬家屋を用いたフィールド試験を行った。家屋外からスピーカで低周波音を出力し、室内の3次元音圧分布計測を行うことによりその低減性能を調べた。その結果、吸音デバイスの設置により室内の音圧分布が大きく変化し、特に部屋の四隅での音圧上昇を抑制できることが明らかになった。また、床面の音圧値の平均低減量は最大7.5dBとなった。

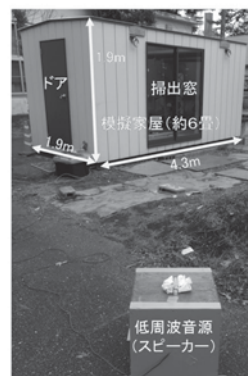


Fig.1: Prefabricated house and loud-speaker



Fig.2: Design of folded resonators



Fig.2: Installation of resonators in the house

1-Q-29

1-Q-29 小型加振器で振動させた窓を制御音源として用いた室内低周波音のアクティブ制御

Active control for low-frequency sound in a room using a window vibrated by a small exciter as a control sound source

◎岩永景一郎, 土肥哲也 (小林理研)

- ◆低周波音50Hzを室内でアクティブ制御(図1, 2)
- ◆加振器で窓を振動させて制御スピーカとして利用(図2)
- ◆制御点(室内隅)周辺1m²に低減効果, 最大14dB(図3)
- ◆低周波音が気になり眠れない場合などへ適用できる可能性

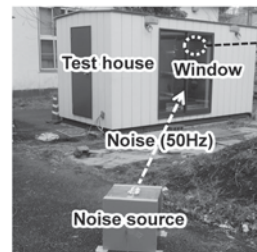


Fig.1: Outside the room (50 Hz noise enters the room)

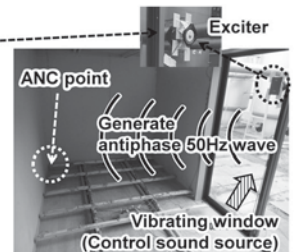


Fig.2: Inside the room (ANC by vibrating window)

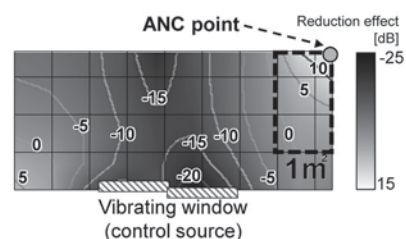


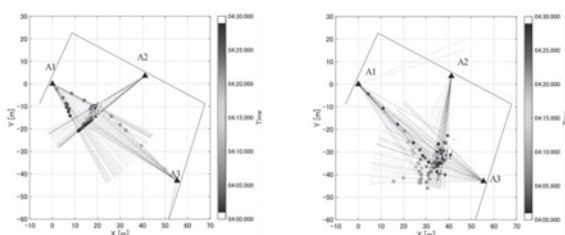
Fig.3: Reduction effect on floor surface

1-Q-30

1-Q-30 複数マイクロホンアレイを用いた移動音源の平面内位置推定に関する実験的検討 その2 壁面で囲われた領域内での検討

Experimental investigation about horizontal source localization of moving noise source using multiple microphone arrays.
- Part 2: Examination in an area surrounded by a wall.
○竹林健一, 星野嗣人, 田中彩, 矢入幹記(鹿島技研)

- ◆建設現場でのリアルタイムでの騒音源位置把握を目的とし、移動する音源の同定について複数マイクロホンアレイの適用を検討している。
- ◆本報では屋外の比較的高い壁に囲まれた領域において、建設重機を稼働させた際の音源位置推定結果について報告を行う。
- ◆Fig. 1 振動ローラとブルドーザの稼働位置推定を示す。
- ◆大まかな位置推定は可能ということが分かったが、アレイの配置精度などによる誤差で推定位置がずれることや、壁面反射による虚音源を推定する場合も生じた。



(a) Vibration roller

(b) Bulldozer

Fig. 1 Results of noise source detection. Left: vibration roller, right: bulldozer
□ : real source position. ○ : estimated source position.

1-Q-32

1-Q-32 離散トークン音声認識におけるドメイン適応の検討

Domain adaptation in discrete token speech recognition.

◎石井敬章, 小松達也, 藤田雄介, 藤田悠哉 (LINE ヤフー)

- ◆離散トークン音声認識を音声データのみでドメイン適応する。
- ◆ドメインに応じて異なる離散トークンの出現パターンを離散トークン系列自身から予測する事で、ドメインミスマッチ軽減を図る。
- ◆ソースドメインに LibriSpeech, ターゲットドメインに TED-LIUM 3 を用いた実験により、有効性を確認した。

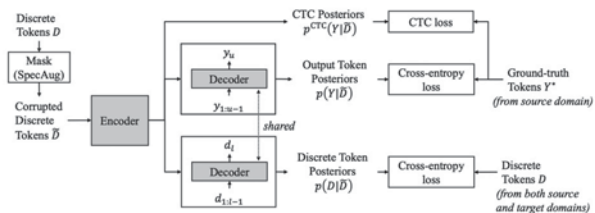


Fig.1: Proposed discrete token prediction loss utilizing unlabeled target domain data.

Table 1: WER (%) results of different models.

Model	Adaptation setup	LibriSpeech		TED-LIUM 3
		clean	other	testset
Baseline		4.0	6.9	10.4
Proposed	Audio	4.0	6.8	9.7
Fine-tuning	Audio and Text	4.5	7.6	8.5

1-Q-31

1-Q-31 歯科治療時の歯導音と気導音

Tooth-conducted and air-conducted sounds during dental treatment

○山田朋美, △阪本 敬, △村上旬平, △林 美加子(大阪大・歯)
△野崎一徳(大阪大・歯病), 桑野園子(大阪大)

- ◆私たちは、自分自身が食べ物を食べる音、歯ぎしりの音、治療中に歯を削る音などを、耳からだけではなく、自分の歯を伝わる振動を音として知覚している。
- ◆音響用語では、頭蓋骨と軟部組織の機械振動を通して内耳へ伝えられることを「骨導」と規定している (JIS Z8106)。
- ◆歯という硬組織を介して上下顎骨を伝わり知覚される音を我々は「骨導音」と区別して「歯導音 (はどうおん)」と呼ぶ。
- ◆本研究では、歯導音計測手法を用いて歯科治療を行った時の歯導音と気導音の計測を行ったので報告する。

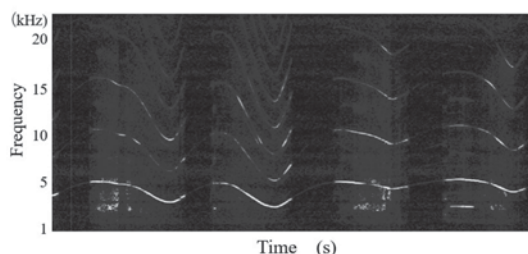


Fig.1: Spectra of tooth-conducted dental drill noise.

1-Q-33

1-Q-33 トピックモデルを用いた教師なし学習によるHuBERTの意味表現向上

HuBERTopic: Enhancing Semantic Representation of HuBERT through Self-Supervision Utilizing Topic Model

◎前角高史 (LINE ヤフー), Jiatong Shi(カーネギーメロン大),
Xuankai Chang (カーネギーメロン大), 藤田悠哉 (LINE ヤフー),
渡部晋治 (カーネギーメロン大)

- ◆代表的な自己教師あり学習モデルである HuBERT は、マスク予測により局所的な文脈情報を学習することで、様々な下流タスクで有効性が確認されている。一方で我々の先行研究では、HuBERT の疑似ラベル系列にトピックモデルを適用し、会話音声クラスタリングできることを示した。これより、疑似ラベル系列から大域的な文脈情報も教師なしで取得可能と期待される。
- ◆そこで本研究では、トピックモデルを用いて HuBERT の意味表現を向上させる手法を提案する。疑似ラベル系列に LDA を適用して得られた発話毎のトピックを推定するタスクを設けることにより、教師なしで HuBERT に対して大域的な文脈情報を与えることが可能となる。実験より、音声認識や SUPERB タスクにおいてベースラインと同等以上の性能を達成した。また、トピックラベルが性別や話者、会話のテーマといった情報を含むことを確認した。

	PR(↓)	ER(↑)	IC(↑)	SID(↑)	SD(↓)	SF(↑)	KS(↑)	SE(↑)
HuBERT	13.89	60.24	88.72	60.48	8.86	80.62	94.22	2.48
提案法	12.97	60.92	90.64	61.82	8.59	81.05	94.87	2.50

Table 1: SUPERB evaluation results.

1-Q-34

1-Q-34 深層学習による口唇情報を用いた音声認識の 多様なノイズタイプにおける評価

Evaluation of Speech Recognition with Lip Information
Using Deep Learning in Various Noise Types

☆伊達龍斗, 西村良太(徳島大), 太田健吾(阿南高専), 北岡教英(豊橋技科大)

- ◆NTCD-TIMIT データセットを用いて、雑音環境に頑健なマルチモーダル音声認識モデルの構築を行った。
- ◆テストセットを用いた提案モデルのノイズ環境下における PER (Phone Error Rate)は50.5%となり、同データセットを用いた先行研究と比較して、平均PERが7.34%改善した。
- ◆口唇情報を使用した際に、音声のみを使用した場合と比較して、どの音素で認識率が改善されたかについても評価を行った。
 - 特に、発音する際に口唇の動きが大きい母音音素において、その効果が顕著であることが示された

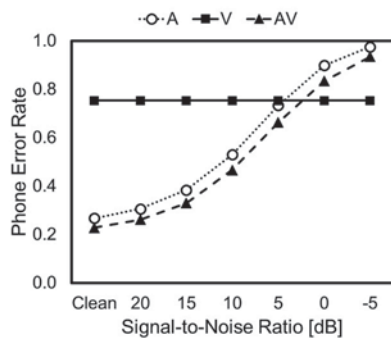


Fig. 1: Phone Error Rate for each SNR

1-Q-36

1-Q-36 Whisper を用いた子ども音声認識における 学年情報の有効性の検証

Effectiveness of using grade information
for Whisper-based child speech recognition.

◎堀井こはる, 俵直弘, 小川厚徳, 荒木章子 (NTT)

- ◆子ども音声は成長によってその特徴が変化するため、汎用的な音声認識モデルでの認識が難しく、特に年少の子どもの音声認識は困難である。
- ◆本研究では最新の汎用音声認識モデル Whisper を子ども音声に適應する際の、子どもの学年情報の有効性を検証した。
- ◆Whisper に対し、全学年の子ども音声で学年非依存の finetuning を行ってから各学年の音声で finetuning を行うことで、話者の学年を考慮したモデルを構築した。
- ◆各モデルについて、各学年の音声で評価を行った結果、評価データの学年にかかわらず特定の学年の音声で finetuning したモデルが全体的に良い精度となった。

Table. 1: Word Error Rates [%] of each model for each grade test data.

		Grade of test data.											
		All	0	1	2	3	4	5	6	7	8	9	10
Grade of finetuning data.	-	17.3	54.1	41.9	22.7	16.7	11.7	14.4	12.6	11.2	10.7	9.5	9.3
	All	15.0	41.6	28.9	20.5	16.6	12.5	11.7	12.7	9.3	9.3	9.2	8.7
	0	15.5	42.3	29.3	21.5	17.3	12.6	12.4	12.8	10.2	9.6	9.7	9.1
	1	14.1	41.1	27.6	19.2	15.6	11.7	11.1	11.4	8.8	8.5	9.0	8.1
	2	14.5	41.7	27.9	19.8	16.1	11.9	11.8	11.9	8.5	9.1	9.0	8.5
	3	14.4	43.1	28.0	19.2	16.2	11.8	11.3	12.0	8.6	9.1	8.9	7.9
	4	15.2	43.1	29.1	19.9	16.5	12.6	12.0	12.8	9.6	9.8	9.2	9.3
	5	16.1	44.2	29.8	21.7	17.1	13.8	13.1	13.5	10.3	10.3	11.4	9.0
	6	15.6	43.6	29.4	20.9	17.0	13.1	12.2	13.6	10.1	10.0	10.4	8.7
	7	15.2	43.5	28.0	20.1	16.5	12.6	12.0	13.1	9.7	10.0	9.6	8.7
	8	16.7	44.1	30.1	22.5	17.8	14.1	13.7	15.0	10.9	10.8	10.9	9.9
9	15.8	43.1	29.4	21.4	17.3	13.2	12.8	13.4	10.2	9.8	10.3	9.2	
10	16.3	44.6	29.8	21.7	17.4	13.9	13.1	14.5	10.5	10.8	10.3	9.7	

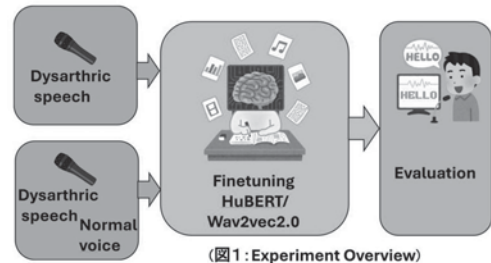
1-Q-35

1-Q-35 構音障害者音声認識における認識精度改善の検討

Improvement of recognition accuracy for dysarthria

○小笠原 朝陽(岩手大) 談宜育(岩手大)

- ◆346万人。日本には構音障害者を含む音声言語障害者がいる。彼らは、綺麗な発音が困難なためコミュニケーションにおける支援が必要となる。
- ◆Google や amazon など世界的企業が開発する音声認識モデルをもってしても彼らの発話を認識することは困難だ。なぜなら彼らの発話には、不透明さ・母音の欠落など健常者とは全く異なる特性を持つからだ。
- ◆Wav2vec2 や HuBERT などの事前学習モデルは、彼らの発話認識に大きく貢献する。彼らの発話をファインチューニングさせることにより、約10%から約40%へと大きく認識精度は向上するのだ。
- ◆社会との壁は物理的だけではない。心理的にも壁は存在するのだ。本研究では、健常者・障害者問わず利用できる音声認識手法を模索することにより、心理的バリアフリーなシステムの実現を目指す。



(図1: Experiment Overview)

1-Q-37

1-Q-37 クロスリンガルベンチマークと類似度分析 による音声 SSL モデルの言語情報分析

Unveiling the Linguistic Capabilities of a Self-Supervised Speech Model
through Cross-Lingual Benchmark and Layer-Wise Similarity Analysis.

◎芦原孝典, デルクロアマーク, 井島勇祐 (日本電信電話株式会社)

- ◆音声 SSL モデルには話者や感情といった情報に加えて、単語や意味といった高度な言語表現も内包している事が示唆されている。
- ◆我々も過去の研究発表会において、SpeechGLUE と呼ばれる General Language Understanding Evaluation (GLUE) に基づいた音声言語理解ベンチマークを通じて、その獲得の有無や程度について発表した。
- ◆本稿では、Japanese GLUE (JGLUE) に基づいた音声言語理解ベンチマーク SpeechJGLUE を構築し、英語・日本語音声 SSL モデルを用いることで、クロスリンガル条件も含めた網羅的な分析を実施した。
- ◆またベンチマーク時に用いた重み付き和の可視化並びに、同一または異なる SSL モデル間の各層の類似度を可視化・分析した。
- ◆実験から、SpeechJGLUE において、一部の英語音声 SSL モデルは日本語音声 SSL モデルよりも高い精度を達成し、言語非依存な言語知識を獲得している可能性が示唆された。

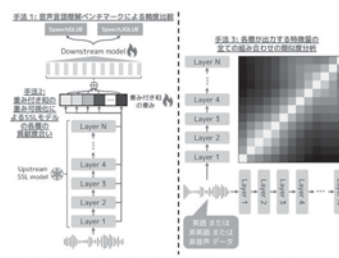


Figure 1: Schematic diagrams of the analysis methods

Table 1: Averaged score in SpeechJGLUE

Upstream	平均スコア
FBANK	42.1
HuBERT LARGE	59.2
HuBERT BASE CSJ	59.4
WavLM LARGE	64.5
BERT LARGE JA	78.7

1-Q-38

1-Q-38 音声コーパス構築のための仮定を追加した発話区間検出法の提案と基礎評価

Proposal and basic evaluation of a voice activity detection with additional assumptions for speech corpus construction.

☆俣野 文義, 小口 純矢, 森勢 将雅(明治大)

- ◆ 音声コーパスの無音区間除去に特化した音声区間検出法を提案
- ◆ VAD にコーパス収録を想定して2つの仮定を追加
 1. 一番最初と最後の音声開始・終了の2点のみを検出
 2. 必ず1つ以上の音声区間が存在
- ◆ SATEN : 始点・終点のみを検出するアルゴリズムを提案
- ◆ ホワイトノイズを重畳した音声で既存VADとの性能比較
- ◆ ノイズを重畳した音声について比較した中で最高性能

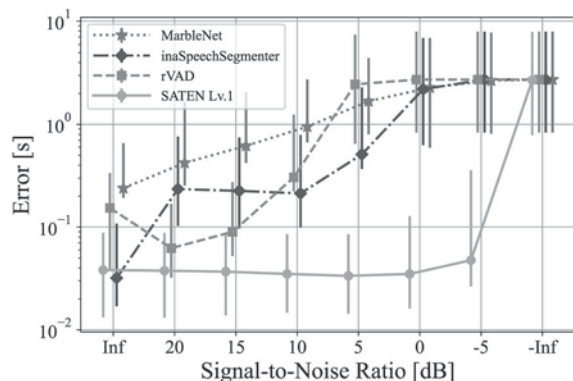


Fig. 1: The horizontal axis shows the SNR of the superimposed noise and the vertical axis shows the absolute value of the error. Plots are median values, error bars are first and third quartiles.

1-Q-40

1-Q-40 FIR型深層波形生成モデルを用いたF0制御可能なEnd-to-End音声合成の検討

F0 controllable end-to-end speech synthesis using deep waveform generation model with FIR filters

○大谷大和, 岡本拓磨 (NICT), 戸田智基 (NICT/名古屋大学), 河井恒 (NICT)

- ◆ 本稿では基本周波数 (F0) の制御が可能な有限インパルス応答に基づくニューラルボコーダ, FIRNet を用いたEnd-to-end 音声合成モデルを提案
- ◆ 従来のEnd-to-end 音声合成モデルと同等以上の性能を示し, 高速な音声合成を実現
- ◆ また F0 制御が可能であることを示したが, 頑健性が不十分のため, 実用に向けてさらなる頑健性向上を検討.

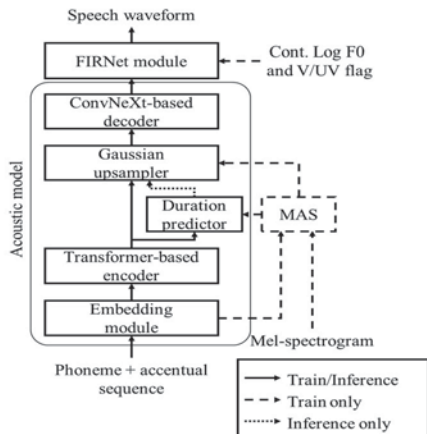


Fig1: Proposed TTS system

1-Q-39

感情演技を伴う日本語オーディオビジュアルデータセットの作成

Creation of Japanese audiovisual dataset with emotional acting

☆廣瀬瑛絵, 大谷祐人, 澤田隼, 大村英史, 桂田浩一(東京理科大・情)

- ◆ 本稿では, 感情演技を伴うオーディオビジュアルデータセットを提案する. 機械読唇, 感情音声合成を行うためには感情演技を含むオーディオビジュアルデータセットが必要不可欠であるが, 既存のデータセットには日本語のものが少なく, プロによる発話でないもの, 話者の性別に偏りがあるもの, 発話内容がバランスでないもの, 1人当たりの文章数が少なく特定話者の学習が困難であるもの, 感情演技を含まないものが多く存在し, これらは機械読唇, 感情音声合成を行うにあたって十分なデータセットとは言えない.
- ◆ そこで, 本研究では8人の役者がITAコーパス原稿(424文)をそれぞれエクマン理論に基づいた7感情(ニュートラル, 幸福, 嫌悪, 悲しみ, 恐怖, 怒り, 驚き)で読み上げた音声と, 正面・横の2面角から撮影した顔の動画を収録した感情オーディオビジュアルデータセットを作成した. 本データセットは, 機械読唇, 表情分析, 感情音声合成, クロスモーダル情報処理など, 様々な研究分野への応用が期待される. なお, 本データセットは研究用途に限り公開する予定である.

1-Q-41

1-Q-41 音声対話システムのためのユーザとの親密さに基づく合成音声制御に関する検討

A Study on Synthesized Speech Control Based on User Intimacy for Spoken Dialogue Systems

☆川西翔貴, 伊藤彰則(東北大), 千葉祐弥(NTT), 能勢隆(東北大)

- ◆ より自然な韻律で会話する音声対話システムのために, 親しみの度合いで音声の特徴を制御する音声合成モデルを構築する. 親しみは高低の二値で与えられるものとし, 親しみに応じた発話文と二値のラベルを組にして音声合成モデルの入力に用いる.
- ◆ 学習には1対1対話を収録したSMOCコーパス [Yamazaki+, '21]を用い, 親しみ低い発話と親しみ高い発話をそれぞれ抽出した.
- ◆ SMOCコーパスのテストセットと学習したモデルの生成を比較したところ, テストセットと同じ条件で生成した場合の方がRMSEが低く, モデルが特徴を正しく学習できていることが示唆された.
- ◆ 全体的な傾向を見ると, 話速やF0の平均はテストセットと同様の傾向であったが, F0のレンジは逆の傾向を示した(Fig. 1). 従ってF0については改善の余地があり, 主観評価への影響を確かめる必要がある.

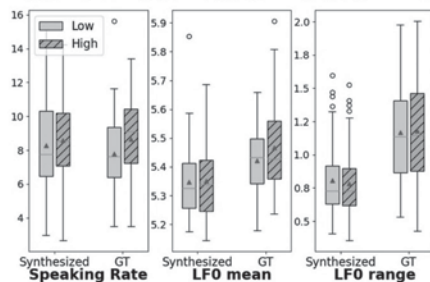


Fig. 1: Acoustic features of synthesized data and ground truth

1-Q-42

明暗の対極的な感情 2 種のコーパスを用いた感情間モーフィングが可能なテキスト音声合成に関する検討

A study on text-to-speech synthesis capable of morphing between emotions using a corpus of two opposing emotions of brightness and darkness.

☆福澤貴裕, 坂野秀樹, 旭健作 (名城大院)

- ◆近年、表現力豊かな合成音声の需要が高まってきているが、音声合成で主に研究されているのは喜怒哀楽といった基本感情で、その他の感情についてはあまり検討がされていない
- ◆本研究では憂鬱の状態を表す「Gloomy」と、驚愕の状態を表す「Astonished」の2種類を用いて、音声合成を行う
- ◆音声合成は Tacotron2 を使用し、ATR 音素/バランス 503 文を用いて平静音声、「Gloomy」、「Astonished」の感情音声を収録した音声を学習させた
- ◆各合成音声はそれぞれの感情音声の特徴に合わせた発声ができおり、原音声との基本周波数の平均と標準偏差を比較しても原音声に近い値で合成できていた
- ◆今後はさらに学習データを増やしていく予定

Table 1: The mean and the standard deviation of the fundamental frequency of the original and synthesized voices for each emotion

音声種類	平均 (Hz)	標準偏差 (Hz)
平静音声(原/合成)	127.5/128.4	28.9/31.7
Gloomy(原/合成)	120.2/121.7	18.6/18.5
Astonished(原/合成)	172.2/178.7	44.3/40.9

1-Q-44

講演取消

1-Q-43

1-Q-43 音声 SSL モデルを活用した動画音声合成

Video Speech Synthesis Using Speech SSL Model.

☆南汰翼, 鏡木時彦 (九州大)

- ◆本研究では、喉頭癌による声帯の摘出が原因で元通りに発声できなくなった場合における新たな代用音声として、深層学習を用いた口唇動画からの音声合成を検討する。
- ◆近年、動画音声合成の分野では AVHuBERT などの自己教師あり学習 (Self-Supervised Learning: SSL) モデルの有効性が示されており、これを基にした音声の高精度な合成が可能となっている。しかし、自然音声に近い品質の合成音声はまだ実現されていない。
- ◆本研究では、AVHuBERT をベースとした動画音声合成モデルの推定残差を小さくするために、音声自体の文脈を学習した SSL モデルである HuBERT を利用した新たなネットワークを提案する。
- ◆提案手法は二段階で構成される。まず、AVHuBERT をベースとしたモデルにより、動画からメルスペクトログラムと HuBERT 中間特徴量を推定する。次に、推定された HuBERT 中間特徴量を入力として、HuBERT の Transformer 層を学習させる。HuBERT 中間特徴量は自己教師あり学習においてマスク対象となる特徴量であるため、その後の Transformer 層を再学習することで推定残差の軽減を狙った。
- ◆実験では、男女二人ずつのデータセットを用い、PESQ, STOI, ESTOI, WER の客観評価指標を基に評価を行った。結果として、HuBERT を用いることで PESQ, STOI, ESTOI の改善を確認した。
- ◆今後は、主観評価実験を通じてさらなる性能評価を行う予定である。

1-Q-45

1-Q-45 低次元音響パラメータを用いた End-to-End ニューラルフォルマント合成の検討

End-to-End Neural Formant Synthesis Using Low-Dimensional Acoustic Parameters

☆小林清流, 小坂哲夫 (山形大), 能勢隆 (東北大)

- ◆一般に、ニューラルボコーダは高次元音響特徴量を用いて収録音声に忠実な音声を合成することが可能
- ◆しかし、ピッチやフォルマントなどのより低次元の音響パラメータでは制御できず、合成時に音韻や声質の制御が困難
- ◆本研究では、F0 やフォルマントなど 7-9 次元の音響パラメータから音声波形を合成する End-to-End 深層モデルを提案
- ◆結果として HiFi-GAN による分析合成や End-to-End 構造を有しない従来法 (neural formant synthesis : NF) と比較して同等以上の性能を達成

Table 1: Results of objective and subjective evaluations for analysis synthesis

Model	UTMOS	VUV [%]	RMSE [cent]
HiFi-GAN	2.81	3.66	43.10
NF (baseline)	2.90	4.11	39.57
E2E-NF	2.64	4.10	38.24
E2E-NF+	2.97	3.93	42.76
E2E-SiFi-NF	3.10	3.90	28.76

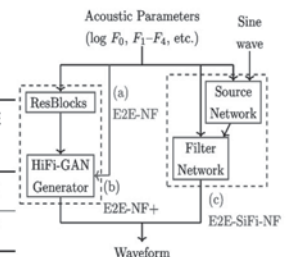


Fig. 1: Architecture of proposed models. (a) E2E-NF, (b) E2E-NF+, (c) E2E-SiFi-NF

1-Q-46

1-Q-46 キャラクター画像からの音声合成のための 音声・顔画像ペアデータセット構築に関する検討 A Study on Construction of Speech and Face Image Pair Dataset from Character Images for Speech Synthesis

☆菊池遥斗, 能勢隆, 伊藤彰則(東北大)

- ◆本稿では、キャラクター画像からそれに適した音声を生成する手法の実現を目指し、キャラクター顔画像・音声・テキスト書き起こしからなるACFV(Anime-Character Face Voice)データセットの構築手法の提案と、構築したデータセットの分析を行った。
- ◆ゲームから抽出した音声に対応する各キャラクターの顔画像を抽出し、類似した顔画像をキャラクター毎にDanbooru データセットから複数選択し、バリエーションのあるデータセットを構築した。
- ◆t-SNEによる顔特徴ベクトルの分布を分析した結果、提案法により元キャラクターに類似した多数の音声・顔画像ペアを作成できていることを確認した。

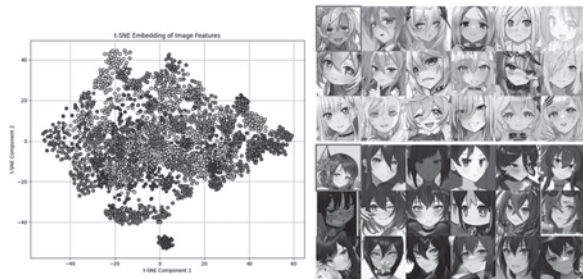


Fig.1: Character face image features visualized by t-SNE

Fig.2: Example of collection of similar character face images

1-Q-48

1-Q-48 吃音者音声合成のための 音声認識を用いたテキストラベル修正 Text Label Modification Using Speech Recognition for Speech Synthesis of Stuttered Speech

☆長久保 諒¹, 山下 陽生¹, 高島 遼一^{1,2}, 安井 美玲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、吃音者の声質でかつ流暢なテキスト音声合成システムの作成を目的として、音声認識を用いてテキストラベルの修正を行う手法を提案する。
- ◆吃音者音声では、吃音症状によって連続発話中に休止が発生し、音声とテキスト間に不整合が生じる問題がある。
- ◆提案手法では、吃音者音声から音声認識モデルと言語モデル(読点のランダム挿入を行ったテキストで学習)によって音声認識を行い、認識結果の読点を音声合成モデルの学習に用いるテキストに挿入し不整合の解消を目指す。

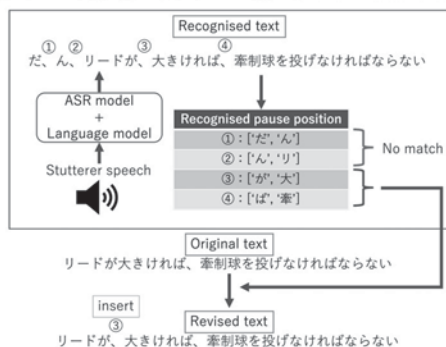


Fig.1 Overview of the proposed method

1-Q-47

1-Q-47 少量な不明瞭音声を用いた適応による 舌垂全摘出者の話者性を持つ テキスト音声合成の検討

Text-to-speech synthesis for glossectomy patients by speaker adaptation using small amounts of glossectomy patient's speech

☆岡村優頼(岡山大学院・HS 統合科学研), 原直, 阿部匡伸(岡山大)

背景・研究目的

- ◆舌垂全摘出者: 舌を半分以上切除した患者, 発音が不明瞭
- ◆健常者の音声で学習された複数話者 TTS モデルを不明瞭な音声で話者適応すると, 話者類似性と音韻明瞭度がトレードオフになりやすい。
- ◆患者の負担を考慮すると, 大規模な収録を行うことは難しい。
- ◆本研究の目的は, 少量で不明瞭な舌垂全摘出者の音声を用いて, 明瞭で話者類似性の高い目標話者の音声を合成することにある。

提案方式

- ◆舌垂全摘出者と健常者で異なる音素ラベルを使用し,
- ◆舌垂全摘出音声と他人の健常音声を混ぜたデータセットで適応。
- ◆目標話者+JVS1 文 100 話者条件で明瞭度と話者類似性がともに改善

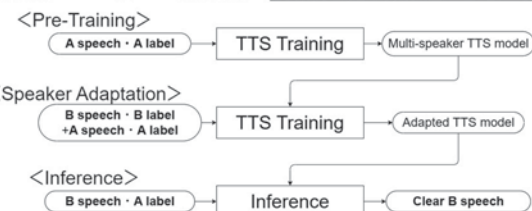


Fig.2: Training and Inference Processes. (A: Healthy patients, B: Glossectomy patients)

1-Q-49

1-Q-49 指示文によって感情変化を制御可能な TTS システムの検討 A Study of a TTS System Capable of Controlling Emotional Changes through Prompts

☆山本麟太郎, △中山英樹(東大院・情報理工)

- ◆テキスト音声合成における感情表現の制御性向上, 特に感情変化の表現を目的として, 指示文ベースの発話スタイル制御で用いることができる新たなアノテーション生成手法を提案する。
- ◆アーキテクチャの概略図を Fig. 1 に示す。オーディオブック読み上げコーパスの文章を LLM によって再帰的に分析し, 文脈から推測した当該文章の感情あるいは感情変化を既存のスタイル音声合成システムに入力する。
- ◆提案手法による合成結果を, 従来の TTS, また一発話につき一つの感情のみに対応させた場合(Baseline)と比較した。音声品質の指標である UT MOS について品質の低下が見られた一方で, 感情表現の正解率では提案手法が Baseline を上回っている。このことから, LLM には単一の感情だけではなく感情の変化までも推測する力があると言える。

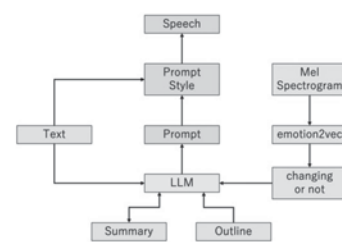


Fig. 1 The architecture of generating both prompt and speech.

Table 1 Results on naturalness and emotional expressiveness of speech.

	UTMOS (↑)	Accuracy (↑)
VITS	4.09	-
Baseline	3.71	0.383
Proposed	3.65	0.612

1-R-1

1-R-1

凸最適化に基づく劣決定BSSに対するADMMアルゴリズムの適用

Application of ADMM algorithm to convex-optimization-based underdetermined BSS

☆皆川朋樹(農工大), 松本和樹(早大), 山田宏樹, 矢田部浩平(農工大)

従来手法 空間モデルを事前に固定することで劣決定音源分離を凸最適化問題として定式化

- 主双対近接分離法(PDS)を用いており空間モデルの設定によっては最適化が遅い

提案手法 交互方向乗数法(ADMM)に基づく最適化アルゴリズムを導出

- 条件によらず安定した最適化を実現
- PDSよりも高い分離性能を達成

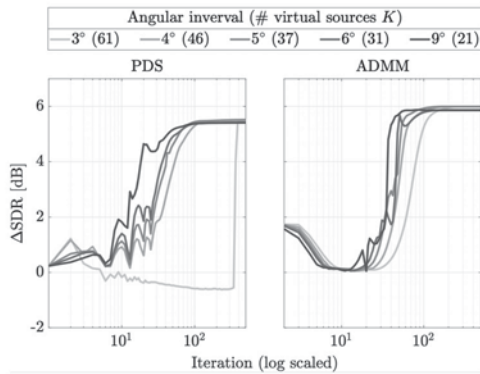


Fig. 1 Comparison of separation performance per iteration

1-R-3

1-R-3 光レーザマイクロホンを用いたラフピント収録音声の深層複素畳み込み再帰型ネットワークに基づく音声強調の評価

Evaluation in speech enhancement based on deep complex convolution recurrent network for recorded speech with rough-focused optical laser microphone

○中野裕貴(立命館大院), 秋庭庭, 岩居健太, 西浦敬信(立命館大)

光レーザマイクロホンは、音により振動している物体に対してレーザ光を照射し、音の計測を行うシステムである。ラフピント収録、すなわちレーザ光が非焦点状態のまま収録を行うことで、焦点距離の変化に頑健な収録が可能となる。この場合、高周波成分の減衰や突発性雑音の混入により、収録音声の音質が劣化する。この問題に対し、深層学習に基づく振幅スペクトルの復元による音声強調が提案されているが、この手法では、位相成分の考慮が行われておらず、音声強調性能が不十分であると考えられる。本稿では、深層学習に基づく複素スペクトルの復元による音声強調を提案し、新しい損失関数について検討した (Fig. 1)。評価実験の結果、提案手法による収録音声の位相群遅延の復元 (Fig. 2) と、一部条件で提案手法の有効性を確認した。

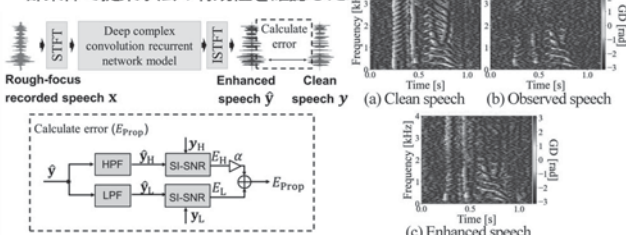


Fig. 1 提案手法の概要

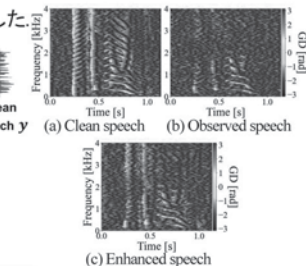


Fig. 2 位相群遅延スペクトログラム

1-R-2

1-R-2 畳み込み型非負値行列因子分解に基づく音響的变化の可視化手法の基礎検討

Basic study of visualization for acoustic variation based on convolutive nonnegative matrix factorization.

☆大村美結, 大島風雅, 中山仁史(広島市大院)

- ◆非負値行列因子分解 (nonnegative matrix factorization: NMF) の表現力に着目し、音素を音源とした音声分解を行うことで音素ごとの新しい音響的特徴の可視化を検討してきた。
- ◆音響的特徴の可視化に伴い従来の NMF では必要としない基底行列の正規化処理を加えた Euclidean metric regulated (erNMF) の提案を行った。
- ◆本稿では NMF による音声分解の分離性能を向上させることを目的として NMF の基底関数の時間的表現が可能な Convolutional NMF の導入を行い、正規化処理を加えた新たな erConvolutional NMF を提案する。
- ◆実験結果より、提案手法は従来手法と比較して推定信号における各基底の大小関係について安定した表現を得ること確認した。

Table 1: Amplitude ratio of pattern 2 when pattern 1 is referenced.

	Ref.amp.	Conv.NMF	erConv.NMF
Pattern1	1.00	1.00	1.00
Pattern2	2.00	0.53	2.99

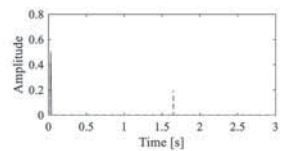


Fig. 1: Reference activation.

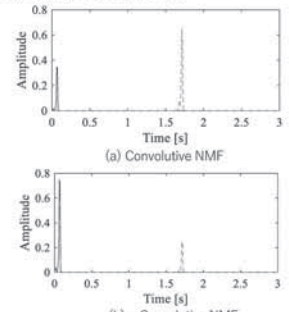


Fig. 2: Visualization results of each Activation.

1-R-4

1-R-4

劣勾配射影に基づく適応的更新を用いた位相復元

Audio phase retrieval with adaptive updates based on subgradient projection

◎ 赤石夏輝, 山田宏樹, 矢田部浩平(農工大)

位相復元: 与えられた振幅スペクトログラムのみから位相の情報を復元する技術

従来 交互方向乗数法を用いた手法 (ADMM-PR) は少ない反復回数で良い性能を示す

課題 振幅が劣化している場合には更新が不安定になる → 更新度合いの調整が性能向上につながる

提案 劣勾配射影によって実現できる適応的更新を ADMM-PR に組み込む

結果 反復初期の更新が安定化 → PESQ・STOI が向上

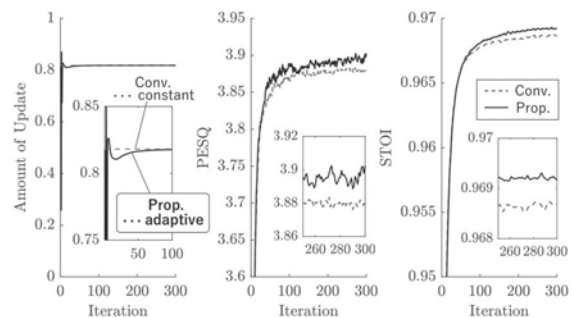


Fig. 1 Comparison between conventional and proposed method

1-R-5

1-R-5 非負値行列因子分解を用いた 深煎り珈琲焙煎音の音響モデリング

Acoustical modeling for dark roast coffee sound
using Non-negative Matrix Factorization

☆向井健悟 (高松桜井高), 大島風雅, 中山仁史 (広島市大院)

- ◆本論文では先の報告で割愛した NMF を用いた周波数計測の着想と、深煎り焙煎音を対象とした時間的変化における周波数特性を示す。
- ◆一般的な NMF では基底関数とアクティベーションの演算結果として信号強度が担保されるため、周波数計測技術として直接利用できることができない。
- ◆ユークリッドノルムで信号強度を正規化した erNMF (Euclidean Metric Regulated NMF) を用いた音響モデリングを試みた。
- ◆焙煎条件に関わりなく同様の音響的変化の傾向が確認され、2ハゼ以降で音響的変化の分岐点とする挙動を確認した。(Fig.1 及び Fig.2)

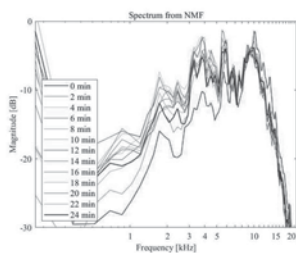


Fig.1: Dark roasting by frying pan

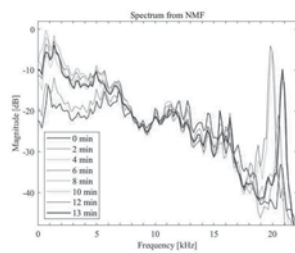


Fig.2: Dark roasting by Coffee Roast Machine

1-R-7

1-R-7 頭部装着マイクロフォンアレイにおける 自己音声の相対伝達関数計測

Relative transfer function measurement for self-voice
in prototype of head-mounted microphone array

☆風間香伽, 中嶋大志, 小野順貴 (都立大)

- ◆背景
 - 補聴器などの頭部装着マイクロホンアレイにおいて、自己音声の相対伝達関数は、自己音声のキャンセル回避等に重要
- ◆課題
 - 一般的な相対伝達関数計測では、白色雑音や TSP 信号などの広帯域な既知の音源信号を用いる
 - 自己音声の相対伝達関数計測では、音源は音声であり、前述のような信号を用いることはできない
- ◆概要
 - 自己音声の相対伝達関数計測に適切な発声手法と、計測した相対伝達関数の個人性について評価
 - 独自に試作したメガネ型マイクロホンアレイを用いて実験

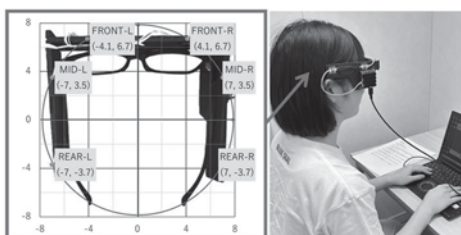


Fig.1: Recording with a glasses-type microphone array

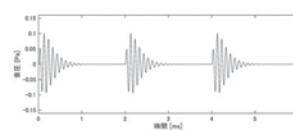
1-R-6

1-R-6 電磁界ばく露により生じる繰り返しパルス音 のモデル化とその聴覚特性解析

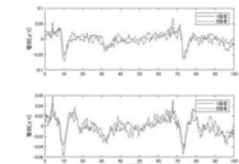
Modeling and analysis of repetitive pulse sound caused by exposure to
electromagnetic fields

☆ハディ平和, 大久保寛(都立大), △寺尾 安生(杏林大), 多氣 昌生(都立大), 鈴木 敬久(都立大), △岸岡 翔馬(都立大), △岡本 真由美(杏林大), △湯浅 薫(杏林大)

- ◆マイクロ波聴覚効果として知られる電磁界ばく露現象があるが、その発生機序は完全には解明されておらず、ばく露評価の観点からマイクロ波を人体に照射する実験には制限が多いことから、この現象で聞こえるとされる音に関する知見は少ない。
- ◆本研究ではマイクロ波聴覚効果で聞こえる音を繰り返しパルス音により再現し、最小可聴値の計測と誘発電位計測により、その聴覚特性の解析を試みた。
- ◆極限法を用いた最小可聴値の測定実験では、純粋なサイン波よりも繰り返しパルス音は聞き取りづらいことがわかった。
- ◆誘発電位計測では、提示する音圧を上げていくことでクリック音と同様に聴性中間潜時反応 (MLR) が観測できることがわかった。



Repetitive pulse with a repetition frequency of 500 Hz



MLR of subject A

1-R-8

1-R-8 24ch スピーカアレイを用いた 生活環境音の最適再生法の一検討

Study of the optimal reproduction method for living environment sounds
using 24ch speaker array.

☆岡田拓真, 山高正烈 (愛知工科大)

- ◆本研究では、Fig.1 に示す自作した 24ch 音空間再生システムによる聴取環境を想定し、Ambisonics 録音方式によって収録した生活環境音を再生する場合の最適な再生方法の検討を行った。具体的には、Ambisonics マイク (AMBEVO VR MIC, SENHEISER) を使用して喫茶店における生活環境音 (コーヒーを淹れる音と音楽, 人の声等) を収録し、収録した音信号を複数の再生方式に変換させ、最適な再生方法について検討した。
- ◆再生方式は、水平面 8ch および 24ch のモノフォニック方式、2ch ステレオ方式 2 種類 (Figure-Of-Eight, Cardioid) と、水平面 8ch および 24ch の一次アンビソニクス (FOA) 再生方式の、計 6 種類を採用した。
- ◆高次感性評価の結果、モノフォニックの評価が低く、8ch FOA 再生方式が高い評価となった。方向知覚の結果からは、FOA 再生方式が 8ch と 24ch とともに、左右判別がある程度可能であることが示された。



Fig.1: 24ch スピーカアレイ

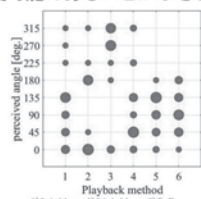


Fig.2: 角度分布図

1-R-9

1-R-9 パニングを用いた 3D 音響信号の伝送・再生に関する検討*

Transmission and binaural rendering of 3D audio signals with panning

○竹内健太, 西口正之, 安倍幸治, 渡邊貴治 (秋田県立大)

- ◆ヘッドトラッキング機能のついたヘッドフォンをPCやスマートフォンにワイヤレス接続した際に、受聴者の頭部運動に伴い必要となる頭部インパルス応答(HRIR)の更新における伝送遅延の影響を低減する方法を検討した。
- ◆パニングを用いたレンダリングを用いることで、ヘッドフォン側でバイノーラル化ができるようにした。
- ◆処理のパラメータが、音源方向とパニングの代表方向の方位差のみに依存するようにし、絶対的な方位の影響を受けにくくした。
- ◆上記の変更を行った結果、頭部回転による SNR の劣化を抑えることができた。

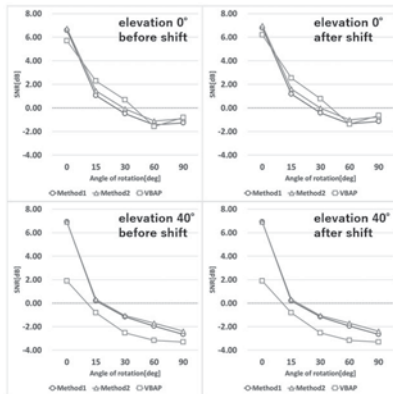


Fig.1: Average of SNR over all azimuth

1-R-11

1-R-11 サイバー・リアル波面整合音響処理 ~Sound SyReal~の実現に向けた第1波面の法則に関する検討

A Study on the Law of the First Wavefront for the Realization of Cyber-Real Wavefront-Matching Acoustic Processing ~Sound SyReal~.

○伊藤弘章, 矢澤櫻子, 渡邊悠希, 千葉大将, 川瀬智子, 加古達也, 磯貝愛, 野口賢一, 中山彰(NTT)

- ◆未来の音環境がリアル音とサイバー音が高度に融合した世界となることを見据え、サイバー・リアル波面整合音響処理~Sound SyReal~というコンセプトを提唱 (1: Synchronous + Cyber + Real を組み合わせた造語)
- ◆オープンイヤー型イヤホン等を活用し、リアルな音波到来と同時にもしくは先行してサイバー音を耳元で重ねて出力し聴覚効果を付与
- ◆第1波面の法則に着目し、スピーカとオープンイヤー型イヤホンの組み合わせ受聴時に、音像が融合して知覚されるかを調査

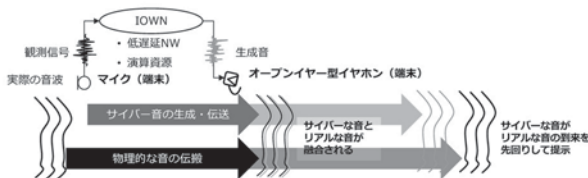


Fig.1: Overview of Sound SyReal.

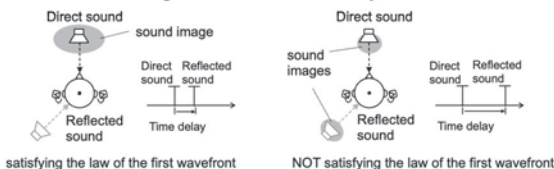


Fig.2: The law of the first wave front and sound image separation.

1-R-10

1-R-10 小型半円形スピーカアレイを用いた耳介近傍音場合成による方向定位

Directional localization of near-ear sound field synthesis with a small half circular loudspeaker array

☆佐久本涉吾, 羽田陽一(電通大)

- ◆小規模な音場合成の実現を目的とし、ヘッドホンのように使える、左右の耳に対応した小型半円形スピーカアレイで耳介周辺のみで音場を合成するシステムの作成と、それを用いた方向定位実験を行った。
- ◆駆動信号の計算には Weighted Pressure-Matching 法を使用し、各二次音源のゲインを制御するために正規化パラメータを行列化した。また、頭部周辺の音圧は、頭部を剛球に近似することで理論的に計算した。
- ◆方向定位実験の結果、提案手法は、KEMER の HRTF を使ったバイノーラル再生と同様の定位傾向を示し、前後誤りと頭内定位はより少なかった。一方で提示角度から 30° ずれた角度に定位しやすいという課題も明らかとなった。

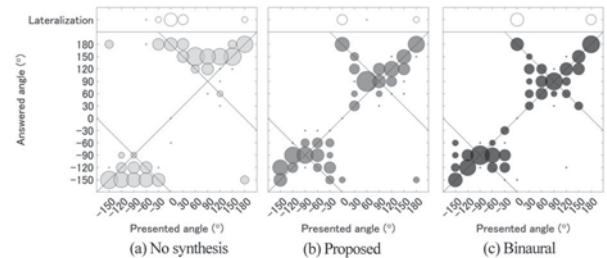


Fig. 1 Bubble chart of angle localization

1-R-12

1-R-12 HOA における音源の幅の提示方法に関する検討

Representation of sound source width using HOA

☆小室奈生, 羽田陽一(電通大)

- ◆HOA において、広がりを持つ一次音源の到方向分布を分析・合成することで音源の幅を提示する手法を提案する。
- ◆42 ch. 球面スピーカアレイを用いて、Fig.1に示すような分布の広がりを変化させる実験と時間変化させない実験を行った。時間変化させる実験では音源の広がり感を一対比較で相対的に、時間変化させない実験では広がり感を絶対的に評価してもらった。
- ◆Fig.1 (c), (e)のように音源の中心付近が抜けた形状の分布の方が、それ以外の分布と比べてより音が広がって聞こえる傾向にあることが分かった。

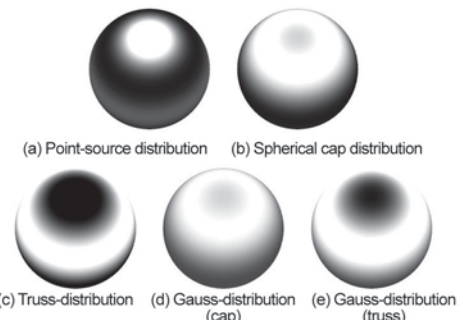


Fig.1: Continuous distribution of primary sound source strength

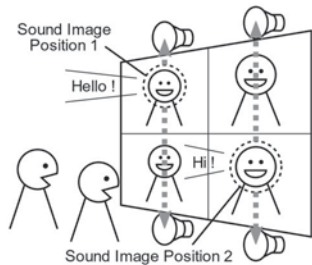
1-R-13

1-R-13 Multiple Vertical Panning を用いた 立体音響システムにおける 移動距離の奥行表現への影響

Effect of Moving Distance on Depth Expression
in 3D Audio System Using Multiple Vertical Panning

○木村敏幸(東北学院大・情報)

- ◆大画面ディスプレイに適合する立体音響システムとして、これまでに Fig. 1 に示すような新たな立体音響システム (Multiple Vertical Panning, MVP 方式) を提案し、実用化の可能性も示してきた。さらに、近年研究されている裸眼立体映像に適合した奥行きを表現するために、音量変化による奥行表現手法についても検討してきた。
- ◆本報告では、これまでに提案してきた MVP 方式による奥行き表現の可能性を検討するために、移動距離を変化させた場合の条件も加えたうえで、映像と音の一致度に基づいた臨場感を評価した。その結果、点音源の物理的な音量変化を模擬すれば最も臨場感が向上するということが分かった。また、MVP 方式がステレオホンックよりも臨場感が向上する場合があることも分かった。



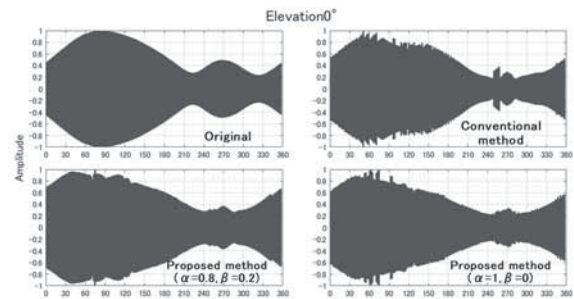
1-R-14

1-R-14 学習によって生成した頭部インパルス応答 を用いたバイノーラルレンダリングにおける コスト関数の検討

Considerations on the Cost Function for Training Head-Related Impulse Responses Used in Binaural Rendering

☆法師人和希, 西口正之, 渡邊貫治, 安倍幸治 (秋田県立大)

- ◆VR や AR のための 3D オーディオ再生において、バイノーラルレンダリングの演算量を削減するために、全地球の頭部インパルス応答(HRIR)を、事前に定めた数個の代表方向 HRIR の合成により模擬する検討を行った。その際、全地球の HRIR を用いて、学習により代表方向 HRIR を生成する方法を検討した。
- ◆前報において歪みの期待値を最小化させる学習を行ったが、必ずしも滑らかな移動音を再現できていないことが確認された。
- ◆本稿では、学習におけるコスト関数を改善することで、移動音の滑らかさを得られないか検討した。
- ◆学習時のコスト関数に制約条件を加えることで Fig.1 の左下のグラフのように右上の従来法よりも滑らかな移動音が作成することができた。



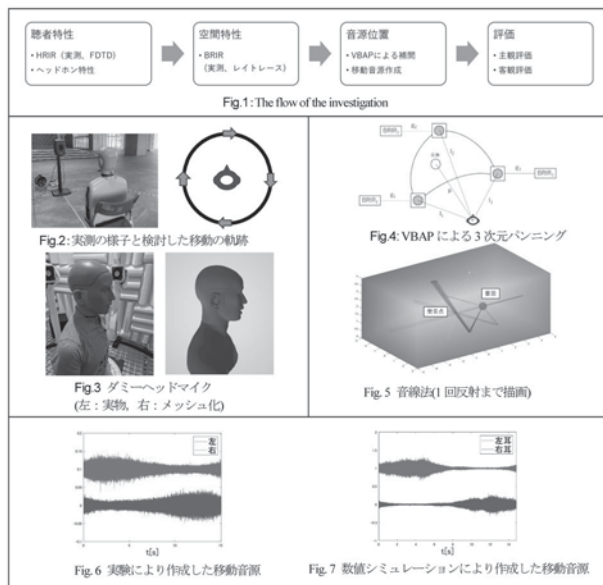
1-R-15

1-R-15 実測と数値シミュレーションを組み合わせた バイノーラル再生のための移動音源の実装

Implementation of Moving Sound Sources for Binaural Reproduction Using a Combination of Measurements and Numerical Simulations

☆高麗瑠太(都立大), 込孫暢, 大久保寛(都立大)

- ◆本研究では、バイノーラル再生のための移動音源の実装方法を検討・提案した。提案法では、ダミーヘッドマイクを用いた実測および数値シミュレーションを組み合わせる手法を検討し、音像の遷移に対する評価を行った。



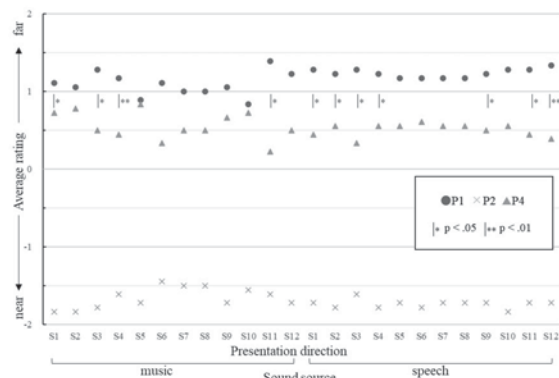
1-R-16

1-R-16 オープンイヤー型イヤホンとスピーカ同時再生 におけるゲイン制御による音像定位・距離評価

Evaluation of sound localization and distance using gain control in simultaneous playback of open-ear headphones and speakers.

○野口 賢一, 千葉 大将, 渡邊 悠希, 小塚 詩穂里, 加古 達也, 伊藤 弘章 (NTT), 亀川 徹 (東京藝術大)

- ◆周囲で発生する音を自然に聞くことが可能であるオープンイヤー型イヤホンの特徴に着目し、イヤホン再生音とスピーカ再生音を同時に聴取することによる音像定位の距離制御を試みる。
- ◆受聴者毎の合成音像の知覚を事前調査し、音源提示方向ごとに設定した再生ゲインを用いることで、提案手法により、イヤホンのみより音像定位精度が向上することを明らかにした。
- ◆また、提案手法により、スピーカと受聴者の間に、音像距離を知覚する傾向があることを示した。(Fig. 1)



1-R-17

1-R-17 極小領域オーディオスポットにおける側帯波分割に基づく音声漏洩の評価

Evaluation of speech-leakage for pin-spot audio based on multiple sideband wave.

☆岩上瑞希(立命館大院), 歌毓庭(立命館大), 中山雅人(阪産大), 西浦敬信(立命館大)

- ◆本研究では、分離放射法による極小領域オーディオスポット形成において、側帯波分割に基づくオーディオスポット外での音声漏洩防止手法を提案する(Fig.1).
- ◆提案手法では、くし型フィルタを用いて側帯波を分割することで、漏洩音を音声として聞き取りづらい信号とすることで、音声の漏洩を低減する。
- ◆本稿では、提案手法を用いて形成した極小領域オーディオスポットと各側帯波用 PAL の放射軸上において、音声の品質に関する評価実験を実施し、提案手法の有効性を確認した。

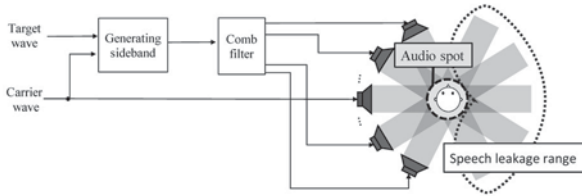


Fig. 1 提案手法の概要図

1-R-19

1-R-19 圧電性シルクセンサの音響特性

Acoustic response of piezoelectric silk sensor

○大久保則男(小林理研, 東理大), 安野功修, 児玉秀和(小林理研), 北島愛菜, 中嶋宇史(東理大)

- ◆シルクはフィブロインとセリシンの2種類の蛋白質から構成されており、不斉炭素を持ち、配向したフィブロイン結晶は D_{∞} の結晶対称性を有する。そのためにシルクを一列に配列し、シート状に成型することで、ずり圧電性を示すことが報告されている。本報告ではシルクシートのずり圧電性を活かしたねじりをずりに変換可能なマイクロホンを作製し、音響特性を評価した。
- ◆評価の結果、この構造による共振周波数が2.6 kHz以下と低く設計できることを確認するとともに、構造によってみかけの圧電定数が向上することを確認した。

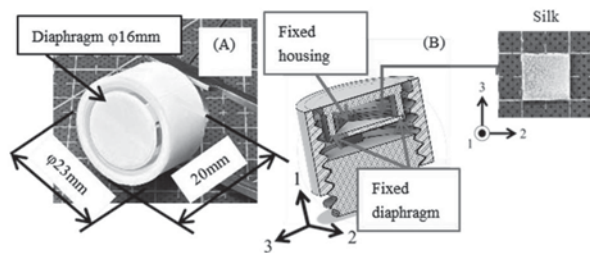


Fig. 1 Structure of piezoelectric silk sensor surface view, (B) cross section view

1-R-18

1-R-18 LiDAR で取得した 3D データへの暗騒音の重畳

Superimposition of background noises on 3D data acquired by LiDAR.

☆水野詩月, 及川靖広(早大理工), 井上敦彦, 寺岡航(早大理工/INSPIRED)

◆背景

音響測定において暗騒音の影響を減らすためにその発生源を特定することは重要である。近年デジタルツインに注目が集まっており、LiDAR を用いた仮想空間上で暗騒音の発生の様子を周波数ごとに表示する手法を提案する。

◆手法

Sound Intensity Scanner AR を用いて暗騒音を計測した結果を、LiDAR で取得した 3D データに重畳し、MATLAB 上で暗騒音の空間音圧分布を周波数ごとに可視化する。結果を動画で表示する。

◆結果

音響計測においてLiDARによる取得したデータの利用を可能とした。エアコン等の暗騒音の発生源が直観的に理解できるようになったほか、音量の小さく発生源が特定しにくい騒音に対してもその発生源を特定できるようになった。

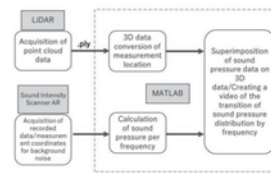


Fig. 1: Measurement procedure

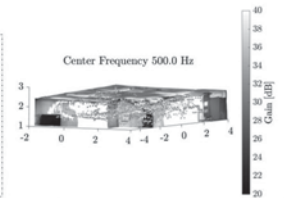


Fig. 2: The sound pressure distribution when the center frequency is 500 Hz

1-R-20

1-R-20 骨伝導音を用いた LMS 法による雑音下の空気伝導音の雑音除去と骨伝導音の音質改善の検討

Noise reduction of air-conducted sound in noisy environments and improvement of sound quality of bone-conducted sound in terms of LMS method using bone-conducted sound.

☆竹内彩華(近畿大院), 上保徹志(近畿大), 中山雅人(阪産大), 折本寿子, 生田顕(県立広島大学), 中迫昇(近畿大)

- ◆情報機器を安全に操作する手段として、音声認識に代表される音声の利用への関心が高まっている。音声の認識率は、静寂環境ではほぼ100%近いが雑音が存在すると認識率が著しく劣化することが知られている。
- ◆そこで雑音対策として、原理的に周囲の騒音が入り難い骨伝導マイクロホンの利用が考えられる。骨伝導音は空気伝導音に比べ音質が劣っていることから、先行研究では、深層学習を用いた音質改善システムの提案やコードブックマッピングによる骨伝導の音質改善が提案されている。
- ◆本研究では、前回の報告に引き続き、空気伝導音の雑音低減の手法を、骨伝導音を入力、雑音下の空気伝導音を出力として、適応信号処理の基礎的な観点から検討し、同時に骨伝導音の音質改善を試みる。
- ◆Fig.1,2 は空気伝導音と骨伝導音の時間波形、Fig.3 は LMS(Least Mean Square)法を適用した出力信号を示す。Fig.3 より雑音が低減されたことが確認できる。

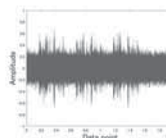


Fig. 1: Air conduction sound.

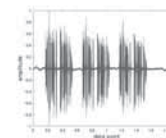


Fig. 2: Bone conduction sound.

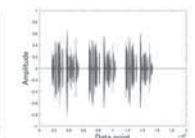


Fig. 3: Output signal.

1-R-21

1-R-21 深層学習モデルを用いた実況動画からのBGM抽出の検討

Development of BGM Extraction from Game Live Video using Deep Learning Model

☆岩本 晃周, 西村 竜一 (和歌山大)

- ◆ゲーム BGM を研究資源として利用できるように整備することを目的に、深層学習モデル Conv-TasNet を用いてゲーム記録からのBGMに重畳された効果音の抑圧モデルの構築及び性能調査
- ◆フリーのBGM80曲及び効果音274種類を使用し、学習用混合音データを作成
 - 異なる3通りのタイミングで効果音が発生するサンプルを用意
 - ✦タイミング: BGMの冒頭から0、2、4秒
 - 混合時のBGMと効果音の音量差を-10dBに調節
- ◆モデルの調査方法として、評価データ7,920サンプルに対してBGM抽出処理前後のSI-SNRの変化量を算出し、BGMと効果音の組み合わせで変化する性能を比較
- ◆BGMの比較では、静かな区間が多い、または周波数帯域が狭いサンプルに対しての抽出が比較的容易であることが示唆された
- ◆効果音の比較では、低音への偏りが大きく、BGMと時間軸上で重複する区間が大きいサンプルに対して抑圧処理が難しくなっていることを確認した

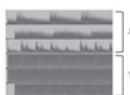


Fig. 1 Examples of BGM signal spectrogram (top 3 and bottom 3 performance)

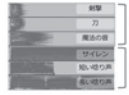


Fig. 2 Examples of sound effect signal spectrogram (top 3 and bottom 3 performance)

1-R-23

1-R-23 4チャンネルマイクロホンアレイの六角形配置による二つの指向性音源の同時定位

Simultaneous localization of two directional sound sources using a hexagonal arrangement of four-channel-microphone-arrays

☆上林大輝, 福田光太郎, 中山雅人, 高橋徹(阪産大院・工学研)

- 【目的】 複数マイクロホンアレイによる非同期音源定位(従来法) 従来手法の計算量削減かつ推定精度改善
- 【手段】 空間を離散化 (Fig.1の■点のみを処理対象に)
- 提案手法1: 空間のビームフォーミングを■点に限定
 - 位置推定方法は従来手法と同様
- 提案手法2: 1) ■点を3マイクロホンアレイでビームフォーミング
2) 出力エネルギーの総和最大化基準で音源定位
3) ■点の中で極大値になる■点を音源と推定
- 【結果】 従来手法、提案手法1、提案手法2の推定精度を比較 (Fig.2)
- 提案手法1: 従来手法と比べ推定誤差の広がりが類似
 - スキャン方向を削減しても推定誤差が同程度
- 提案手法2: 推定誤差0.0mの頻度が高く平均推定誤差が最も低い
 - 計算量削減と高精度化を達成

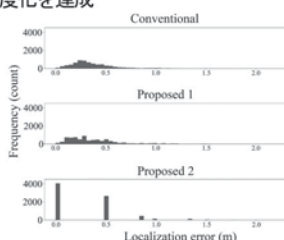
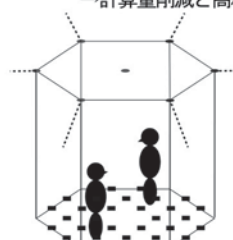


Fig. 1: Localization for two sources. Fig. 2: Histogram of estimation errors.

1-R-22

1-R-22 非負値行列因子分解を用いた韓国単音節の音声分解の基礎検討

A fundamental study of phonetic decomposition of Korean monosyllables using non-negative matrix factorization

☆梨和美佑, 大島風雅, 中山仁史(広島市大院)

- ◆一般的な音声認識では音素単位の音響マッチングが行われている。
- ◆音素は子音および母音の2種類に分けられ、音源や調音構造の違いより各周波数特性は異なる
- ◆音素のアライメントは一般的に重複を考慮せずに時間分割で音声波形が切り出されている。
- ◆音素や音声記号単位で音源は独立していると着想し、従来では考慮されていない音素の重複に注目した。
- ◆各音素を別音源とみなした分離が実現すれば、時系列的なオーバーラップを含む本来収録不可能な音声信号の抽出および解析が期待できる。
- ◆韓国語単音節にNMFを用いることで音素の時間的なオーバーラップと発声のエネルギーの大きさに分解が可能であることを確認した。

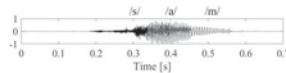


Fig. 1: Waveform of /sam/

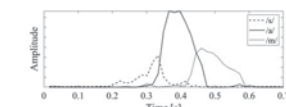


Fig. 3: Each vector of basis matrix

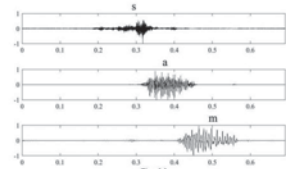


Fig. 2: Speech waveform decomposed by nrNMF

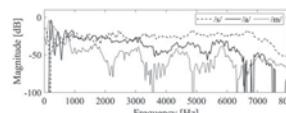


Fig. 4: Each vector of activation matrix

1-R-24

1-R-24 動的なリフタ係数を用いたケプストラム基底分解型独立低ランク行列分析による音声分離

Speech separation by Cepstral-Basis-Decomposed Independent Low-Rank Matrix Analysis using dynamic lifter

☆栗原和真, 大島風雅, 中山仁史 (広島市大院)

- ◆CBD-ILRMA (cepstral-basis-decomposed ILRMA) はILRMAの音源モデルにケプストラム分析を適用したものである。
- ◆CBD-ILRMAでは声帯振動と調音特性に周波数特性を分離することで、初期値依存性による音源分離性能の低下を回避している。
- ◆本稿では CBD-ILRMA の更新回数に対してリフタ係数の値を動的に変化させ、LQ音源モデルにおけるCBD-ILRMAの音源分離性能の向上を試みる。
- ◆動的なリフタ係数を用いることで、CBD-ILRMAの音源分離性能の改善が得られることを確認した。

Table 1: SDR improvements of each dynamic lifter coefficient for proposed CBD-ILRMA using LQ source model.

H	Max	SD	Avg
5 (baseline)	16.91	0.22	16.29
200	18.42	1.16	17.02
300	18.41	1.31	17.04
500	18.39	1.69	16.98
700	18.52	3.24	16.66
800	18.42	3.32	16.52

1-R-25

1-R-25 深層パーミュテーション解決法に基づく ブラインド音源分離の性能評価

Evaluation of blind source separation performance based on deep permutation solver

☆蓮池郁也, 北村大地 (香川高専)

◆目的

- 周波数領域独立成分分析 (FDICA) で生じる周波数毎の分離誤差を考慮した深層パーミュテーション解決法 (deep permutation solver: DPS) の構築

◆提案手法

- 推定誤差量の割合を決める時間周波数行列を作成し, FDICA の推定分離誤差を模倣した学習データを利用

◆実験結果

- Fig. 1 に示すように提案手法 (PS: DPS) が FDICA を適用した信号に対してパーミュテーション問題を解決
- 省サンプルで汎化性能の高いパーミュテーション解決法を実現

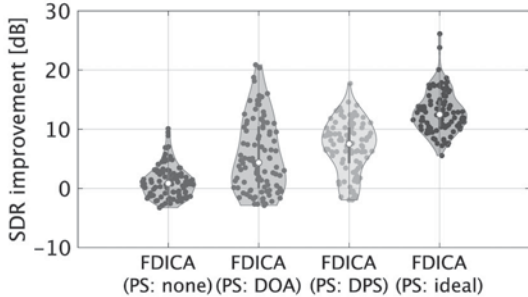


Fig. 1: Violin plot of SDR improvements.

1-R-27

1-R-27 Local equivariance error に基づく サンプリング周波数非依存性評価指標の検討

Investigation on Metric of Evaluating Sampling Frequency Independence Based on Local Equivariance Error

◎今村 奏海(東大/産総研), 中村 友彦(産総研), 高宗 典玄(東大), 矢田部 浩平(農工大), 猿渡 洋(東大)

◆背景: 未学習のサンプリング周波数 (SF) への対応

- SF 非依存 (SFI) 畳み込み層: SF に応じて畳み込み層の重みを生成
- SFI 畳み込み層による SF 適応法は, 入力信号のリサンプリングより一貫した性能を持つことが実験的に確認されている

◆課題: サンプリング周波数非依存 (SFI) 性の評価

- 手法がどの程度 SF に独立か=SFI 性の定量化は行われていない
- Source-to-distortion ratio (SDR) により評価してきたが, SDR は SF に依存しているため, SFI 性の定量化に転用することは難しい

◆提案法: local equivariance error (LEE) に基づく SFI 性評価指標

- LEE:画像処理分野で提案された, 深層ニューラルネットワークの変換に対する頑健性を表す, 同変性を測る指標
- SFI 性をリサンプリングに対する同変性と捉え, 入力の変換としてリサンプリングを用いた LEE に基づく指標を構築

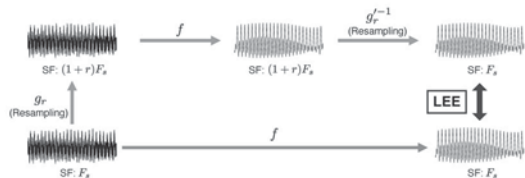


Fig. 1: Schematic illustration of LEE for evaluating sampling frequency independence.

1-R-26

1-R-26 ADMM アルゴリズムを用いた ILRMA における スパース正則化項およびアルゴリズムの比較検討

Comparison of sparse regularization terms and algorithms in ADMM-based ILRMA

◎渡會博子 (農工大), 松本和樹 (早大), 山田宏樹, 矢田部浩平 (農工大)

◆我々は前回, ADMM を用いた ILRMA 型 BSS 手法を提案

- 適切なパラメータ設定下において, 補助関数法に基づく ILRMA と同程度の性能
- しかし, 通常の ILRMA と目的関数が異なっていた

◆本稿では, 通常の ILRMA と目的関数を合わせるためにスパース正則化項を変更した ADMM-ILRMA を導出

- 分離性能の分布の傾向の違いやスパース正則化項が分離に与える影響を確認

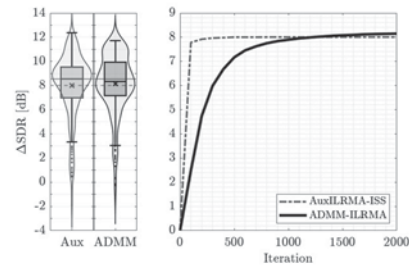


Fig. 1: Comparison of separation performance after 2000 iterations (left) and average separation performance per 100 iterations (right)

1-R-28

1-R-28

1-Lipschitz 畳み込み層を用いた DNN 雑音除去

DNN-based denoising with 1-Lipschitz convolutional layers

☆内田蓮, 山田宏樹, 矢田部浩平 (農工大)

背景

- 1-Lipschitz 連続な DNN
 - ロバストな DNN
 - 最適化アルゴリズムと組み合わせ, 収束性や安定性を保証

1-Lipschitz 連続な DNN のための 1-Lipschitz 連続な畳み込み層がいくつか提案されている

本研究

1-Lipschitz 連続畳み込み層を持つ DNN を用いた雑音除去敵対的サンプルを用いて, ロバスト性を評価

結果

1-Lipschitz 連続な畳み込み層による表現力の低下と敵対的攻撃に対するロバスト性の向上を確認した

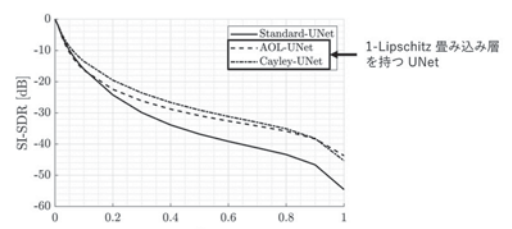


Fig. 1 SI-SDR change in inference results due to adversarial attacks.

1-R-29

1-R-29 低 SN 音声を目的信号に用いる 音声強調学習方法

Training method of speech enhancement using low-SNR speech as the target signal

☆加藤紗耶(龍谷大院・理工学研), 片岡章俊(龍谷大・先端理工)

- ◆スマートフォンやビデオカメラで撮影した雑音が混入した映像から音声のみを抽出し、これを「低 SN 音声」として学習の目的信号に用いた音声強調学習方法を検討した。
- ◆簡単に手に入れられる少量のクリーン音声を利用した学習方法を提案する。少量のクリーン音声のみを目的信号として学習させ、生成されたマスクを用いて低 SN 音声を音声強調処理する。その後、音声強調処理された低 SN 音声を目的信号とし、それに雑音を付加したものをを用いて再度学習を行う。
- ◆低 SN 音声のみで学習を行ったときと、学習に用いたクリーン音声の数が 10, 20, 50, 100 の結果を下表に示す。評価指標は Si-SDR である。
- ◆結果、少量のクリーン音声を用いる学習方法によって評価値が大幅に改善した。しかしクリーン音声の数が多くても必ずしも評価値が高いわけではない。また、同じ SNR の音声であっても雑音の種類によって評価値に大きな差があり、特に人の声の混じった雑音や工場の機械音に対する対策が必要である。

Table: Si-SDR of Experimental Results

Type of Noise	Stationary		Non-Stationary	
	5 dB	10 dB	5 dB	10 dB
Input Speech SNR				
low-SNR Speech	7.33	11.62	7.49	11.72
Number of Clean Speech	10	17.43	19.18	14.35
	20	16.85	18.69	14.23
	50	16.84	19.85	14.66
	100	17.12	19.14	14.27

1-R-31

1-R-31 深層学習を用いた分離した音の変換による 環境音認識の精度向上

Improving the accuracy of environmental sound recognition by transforming separated sounds using deep learning

☆松田祐輔(公立はこだて未来大学)

- ◆環境音認識はあらゆる音を認識して、その場で何が起きているかを把握、予測する技術である。
- ◆環境音認識はそれぞれの音を認識するために音源分離を行い、信号を分離させる必要がある。しかし、分離された信号は他の環境音の影響を受けて、ノイズや欠損が発生してしまい認識率が低下してしまう。
- ◆本研究では、深層学習による生成モデルを用いて、分離された信号を入力とし、似たような音に変換して認識させることを提案し、提案手法を評価した。
- ◆分離した信号と変換した信号を同一認識モデルに入力し、認識率を比較した。結果は特定のクラスでは変換した信号のほうが高くなったが、全体的にみると分離した信号の認識率のほうが高い結果となった。

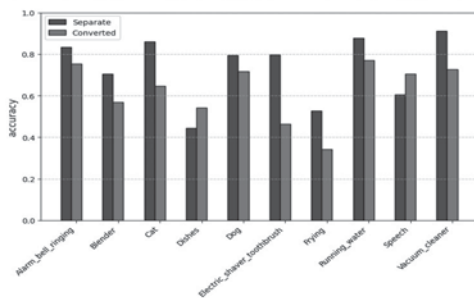


Figure1:分離した信号と変換した信号の認識率

1-R-30

1-R-30 クリーン不要な DNN 音声雑音除去の 実コーパスを用いた性能評価

Performance evaluation of DNN-based speech noise reduction without clean signal using real corpus

☆広政遼汰, 宮崎亮一(徳山高専)

- ◆一般的な音声雑音除去 DNN は、クリーン音声 (clean) を目的信号として用いる Clean-target Training (CTT) で学習しているが、clean は収集難易度が高いという問題がある。
- ◆この問題を解決した手法として、clean に雑音 (noise) が混入した音声 (noisy) を目的信号として用いる Noisy-target Training (NyTT) と呼ばれる手法がある。
- ◆NyTT は無響室などの特殊な環境が不要であり、実環境で収録された noisy (noisy_{real}) を目的信号として用いることができるため、本稿では NyTT と相性の良い noisy_{real} コーパスについて考察し、「LaboroTVSpeech」を使用して実験を行った。
- ◆実験により、目的信号のデータ数を増やすと学習時と評価時で話者のミスマッチがない場合を上回る性能になることがわかった。(Fig.1, 2).

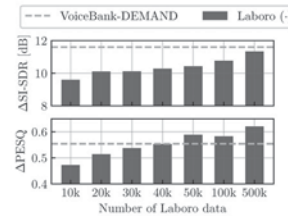


Fig. 1 Improvement in Si-SDR and PESQ when varying number of Laboro.

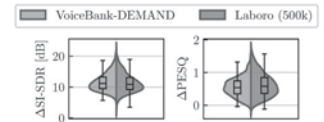


Fig. 2 Scattering of Si-SDR and PESQ improvements in VoiceBank-DEMAND and Laboro (500k).

1-R-32

1-R-32 時間周波数領域での厳密な畳み込みに関する 変換行列の疑似逆

Pseudo-inverse of the matrix related to convolution in the time-frequency domain

☆松山直哉, 山田宏樹, 矢田部浩平(農工大)

- 時間周波数領域での近似畳み込みは、インパルス応答が長いほど精度が悪くなってしまふ
- これまでに、時間周波数領域での厳密な畳み込みを導出した
- 厳密な計算に必要なインパルス応答の時間周波数表現を得るための、順変換行列とその随伴行列を定式化した
- 時間周波数領域での畳み込みを応用するには逆変換が必要

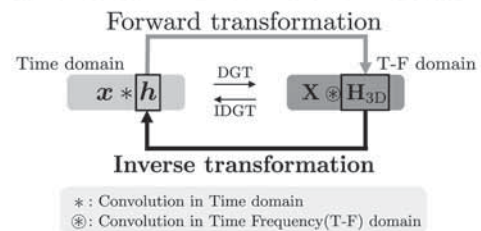


Fig. 1 Overview of the proposed method

- 時間領域のインパルス応答を再構成する疑似逆を導出
- 時間領域と時間周波数領域で一貫した畳み込みが可能

1-R-33

1-R-33 重み付きスパース行列復元に基づく 単チャンネルブラインド残響除去

Single-channel blind dereverberation based on weighted sparse matrix recovery
◎ 饒平名文希, 山田宏樹, 矢田部浩平 (農工大)

目的 単チャンネルの音響信号に含まれる残響を補助情報無しで抑制する単チャンネルブラインド残響除去

以前提案した手法 観測信号を行列リフティングでモデル化
→ スパース行列復元問題として定式化

課題 スパース正則化で用いるグループソフト閾値処理では大きな要素が過剰に抑制されてしまう
→ 各要素のペナルティの大きさを個別に調整する重みを導入した際の性能面での効果を検証

結果 全ての残響時間において性能が向上した

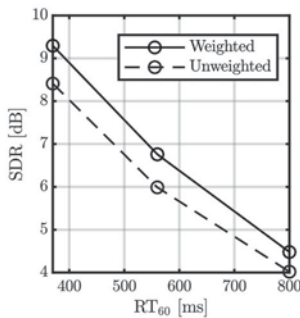


Fig.1 Comparison between unweighted and weighted processing

1-R-35

1-R-35 音響特徴量からの DNN スペクトログラム予測の主観評価

Subjective evaluation of spectrogram prediction from acoustic features based on deep neural network
☆川口翔也 (農工大), 北村大地 (香川高専)

- 我々は前回音高・音色・音量の3つの特徴量から振幅スペクトログラムを予測するDNNデコーダを作成
 - MLP・BiGRU・BiLSTMの3種類のネットワークでそれぞれデコーダを作成
 - 2種類の客観評価指標を用いてBiGRU及びBiLSTMが高精度な予測が可能という結果を得た
- 本稿では前回作成したDNNデコーダを主観評価実験により評価
 - 音楽は人間の感性に関連するものであるため、人間が実際に聞いたうえで評価するべきである
 - DMOS法を用いてBiGRU及びBiLSTMが予測した楽器音が元の楽器音をどの程度近いかを評価 (Fig. 1)

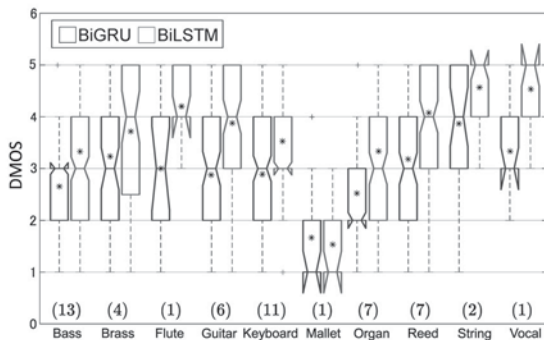


Fig. 1: Results of subjective test with DMOS for each instrument by BiGRU and BiLSTM.

1-R-34

1-R-34

時間周波数領域における 近似畳み込みに関する性質の調査

Investigating characteristics of approximate convolution in the time-frequency domain
☆ 村田佳斗, 矢田部浩平 (農工大), 泉悠斗, 高橋祐, 近藤多伸 (ヤマハ株式会社)

時間領域の畳み込みを
時間周波数領域の畳み込みで近似することができる

問題 時間周波数領域の畳み込みは厳密ではなくその誤差が後段の処理に影響を与える可能性がある

目的 近似誤差について厳密な議論や評価を行うために時間周波数領域の近似畳み込みの性質を調査する

方法 畳み込みに用いる信号の標準基底とそれに対応する係数を用いた畳み込み表現を導出する

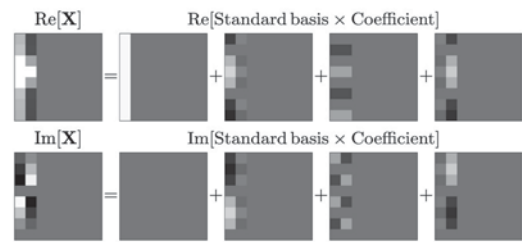


Fig. 1 Expression of the signal (real part and imaginary part) using standard bases and corresponding coefficients

上記の表現を用いた時間周波数領域の
近似畳み込み表現を導出

1-R-36

1-R-36

ℓ_1/ℓ_2 分数ペナルティによるスパース正則化 を用いたAudio Declipping

Audio declipping using sparse regularization with L1/L2 fractional penalty
☆ 角田清香 (農工大), 松本和樹 (早大), 山田宏樹, 矢田部浩平 (農工大)

問題設定: スパース最適化に基づくaudio declipping

- ▶ 振幅超過によりclippingされた部分を復元
- ▶ スパース性を誘導することでclippingにより生じた高調波成分を除去

従来手法: L1ノルムを用いたaudio declipping

- ▶ L1ノルムはスケールに比例し、原点で最小値をとる
 - 全ての成分が小さく見積もられる

提案手法: L1/L2分数ペナルティを用いたaudio declipping

- ▶ L1/L2分数ノルムはスケール不変なスパース正則化項
- ▶ L1ノルムと比べ高い復元性能を実現 (Fig. 1)

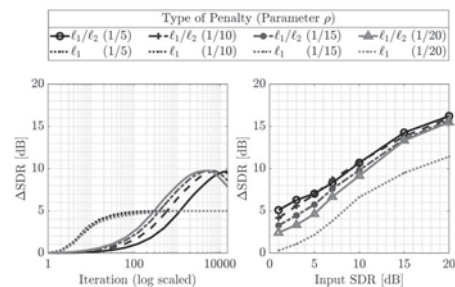


Fig. 1 Performance comparison of conventional and proposed methods

1-R-37

1-R-37 排便前後の腸蠕動音の計測

Measurement of intestinal peristalsis sounds before and after defecation

☆川畑秀斗, 大塚 翔, 中川誠司 (千葉大)

- ◆腹部聴診で検出される蠕動音は腸活動指標の一つとして用いられてきたが、その発生メカニズムには不明な点が多く残る。
- ◆近年、腸音を連続的に観測することで排便時間の予測が試みられている。排泄ケアは介護現場において最も負荷の高い作業の一つであり、排便時間の予測は介護の効率化に大きく寄与する可能性が高い。
- ◆無響室内にて排便前後の腸蠕動音の計測を行い、蠕動音強度と蠕動音発生回数の変化を検証した。
- ◆蠕動音の計測には自作の聴診器マイクロホンを用い、腹部の4箇所計測を行った (Fig. 1)。
- ◆排便後に比べ、排便前の方が蠕動音強度が増加する傾向が認められた ($p=0.0893$, Fig. 2)。一方、蠕動音発生回数には有意な変化は観察されなかった。
- ◆排便前の蠕動音強度の増加は、排便に向けた直腸運動の活性化を捉えた可能性がある。

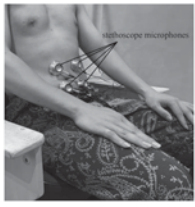


Fig.1 Measurement of peristaltic sound using stethoscope microphones.

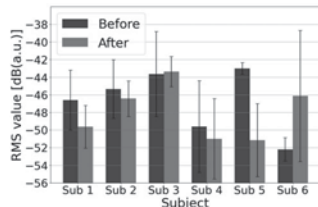


Fig.2 RMS values before (blue) and after (red) the defecation.

1-R-39

1-R-39 Cellular PP マイクロホンを用いた 3次元位置検出

Study on three-dimensional position detection using cellular PP microphone

☆吉井亮介(東理大), 安野功修, 児玉秀和(小林理研), 中嶋宇史(東理大)

- ◆多孔質構造を有する圧電高分子である Cellular Polypropylene(PP)を用いたマイクロホンは、高感度なセンサー材料として期待されている。Cellular PP マイクロホンを用いて音波の指向特性の違いを利用し、**3次元位置検出**の可能性を検討した。
- ◆Cellular PP マイクロホンは**感度周波数特性**を測定した結果、**低周波数領域**では角度によらず**感度がほぼ一定**であり、**30 kHz 以上**では固定角度の違いによって**感度変化が生じた**。
- ◆正十二面体スピーカを用い、音圧の**相対角度依存性**を測定した結果、**低周波領域**では相対角度によらず**音圧がほぼ一定**で、**高周波領域**では**音圧変化が生じる**ことを確認した。
- ◆受信振幅を特徴量、マイクロホンの位置パラメータ r, θ, ϕ を目的変数として、**k 近傍回帰モデル**から決定係数 R^2 を求めた。その結果、 r, θ, ϕ を求める際に最適な周波数を決定することができた。

Table 1 Accuracy of position detection

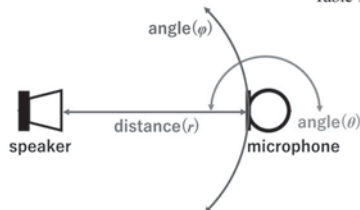


Fig. 1 Relative positions of speakers and sensors

variable	frequency	R^2
r	700 Hz	0.94
	3 kHz	0.88
	10 kHz	0.74
θ	20 kHz	0.94
	60 kHz	0.82
	90 kHz	0.62
ϕ	20 kHz	0.46
	60 kHz	0.74
	90 kHz	0.37

1-R-38

1-R-38 顎顔面形状の違いによる咬合状態を対象とした音響特徴量差に関する研究

A study on acoustic feature differences in vertical occlusal conditions for different maxillofacial geometries

○藤村秀弥 (広島市大院), △村田聡, △堀畑聡, △石井かおり, △武藤佑子, △根岸慎一 (日大・松戸歯), 中山仁史 (広島市大院)

- ◆不正咬合とは顎顔面および歯などが形態、発育および機能に異常をきたす咬合状態のことである。
- ◆不正咬合の治療の目的は咬合と口腔機能の確立並びに不定愁訴のトラブルを解消することである。
- ◆顎顔面形状の違いによる咬合状態を非侵襲的かつ迅速に判断可能な音響分析により評価し、形態的また機能的観点で明らかにする。
- ◆本研究では咬合状態を非侵襲的かつ迅速に判断可能な音響分析による評価を提案する。

	MFCC	Δ MFCC	$\Delta\Delta$ MFCC
1	ns	ns	ns
2	ns	*	*
3	*	ns	ns
4	ns	ns	ns
5	ns	ns	ns
6	**	**	*
7	ns	ns	ns
8	ns	ns	ns
9	ns	ns	ns
10	ns	ns	ns
11	*	*	*
12	*	*	*
13	ns	ns	ns
14	*	*	*
15	ns	ns	ns
16	ns	ns	ns
17	ns	ns	ns
18	ns	ns	ns
19	ns	ns	ns
20	*	*	*

ns: No significant, *: $P < 0.05$, **: $P < 0.01$

Table 1: One-way ANOVA of MFCCs in /ka/.

	MFCC		Δ MFCC		$\Delta\Delta$ MFCC	
	Open	Deep	Open	Deep	Open	Deep
1	ns	ns	ns	ns	ns	ns
2	*	ns	**	ns	**	ns
3	*	ns	*	ns	ns	ns
4	ns	ns	ns	ns	ns	ns
5	ns	ns	ns	ns	ns	ns
6	*	**	**	**	**	**
7	ns	*	ns	ns	ns	ns
8	ns	ns	ns	ns	ns	ns
9	ns	ns	ns	ns	ns	ns
10	ns	ns	ns	ns	ns	ns
11	*	ns	*	ns	*	ns
12	**	ns	**	ns	**	**
13	ns	ns	ns	ns	ns	ns
14	**	ns	*	ns	*	ns
15	ns	ns	ns	ns	ns	ns
16	ns	ns	ns	ns	ns	ns
17	*	ns	*	ns	*	ns
18	ns	ns	ns	ns	ns	ns
19	ns	ns	ns	ns	ns	ns
20	*	ns	*	ns	*	ns

ns: No significant, *: $P < 0.05$, **: $P < 0.01$

Table 2: T-tests of MFCCs in /ka/.

1-R-40

1-R-40 振動板の曲げ波を利用した パネルスピーカの実装事例

Case Study on implementation of a Panel Loudspeaker employed diaphragm bending wave

○河原一彦(九州大・芸工)

- ◆本研究では、パネルスピーカを設計・実装することを試みた。
- ◆設計時には、周波数特性よりも、聴感上の“いわゆる放射の拡散的性質”を重視した。
- ◆本研究で実装した拡散的な特性を重視したパネルスピーカを、曲げ波スピーカ(Bending Wave Loudspeaker: BWL) とよぶこととする。
- ◆本稿では、設計の方針についてのべ、基本特性として、方向別の振幅周波数特性を報告する。

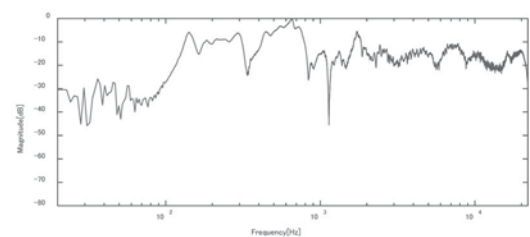


Fig.1: On-axis frequency response of an implemented bending wave loudspeaker.

1-R-41

1-R-41 超低周波音センサの比較校正チャンバの製作
 Making an acoustic small chamber for calibrating infrasound sensors by the comparison method
 ○高橋弘宜, 平野琴, 山田桂輔(産総研)

- ◆超低周波域における計測用マイクロホンの感度校正技術が確立しつつあり、それを基準とした比較校正によって計量トレーサビリティの取れた超低周波音センサの校正も可能となってきた。
- ◆超低周波音センサを比較校正するためのチャンバを製作した(Fig. 1)。20 Hz 以下でチャンパ内の音圧が一様となるよう、チャンパの内寸は 25 cm × 25 cm × 30 cm の直方体にした。
- ◆チャンパ内の音圧が 94 dB となるようにして、1 Hz から 100 Hz までの 7 周波数点でチャンパ内の音圧分布を実測したところ、20 Hz 以下では 0.00 dB ± 0.05 dB 以下で音圧の一様性を確認した。
- ◆試験的に比較校正をした結果、0.4 Hz から 20 Hz までの校正ができることを確認した。校正周波数の下限界が今後の検討課題である。

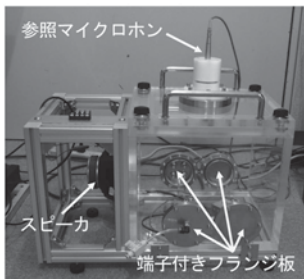


Fig.1: Acoustic chamber designed for calibrating infrasound sensors

1-R-43

1-R-43 M 系列信号を用いた連続音場測定の技術向上に向けた取り組み
 Efforts to Improve the Technology of Continuous Acoustic Measurement Using M-sequence
 ○竹内杏奈, 立花杜斗, △土橋大介(TOA 株式会社), 河原一彦(九州大・芸工)

- ◆従来の音場の拡声の均一性を評価するための音響測定手法には、測定点数の限界と測定に時間を要するという課題がある。
- ◆この課題を解決するため、M 系列信号を用いた連続音場測定法を提案しており、これまで直線経路での測定と精度検証を実施してきた。
- ◆本研究では、実現場のより複雑な条件を想定し、蛇行した経路において測定と精度検証を行った。
- ◆測定点ごとにマイクを配置する従来法の測定結果と、マイクを持ち移動しながら録音する連続音場測定法による測定結果を比較した。
- ◆結果より、従来法と連続音場測定法で得られた測定結果は近似することが分かった。よって連続音場測定法は、直線だけでなく蛇行経路での測定にも適用可能であると考えられる。

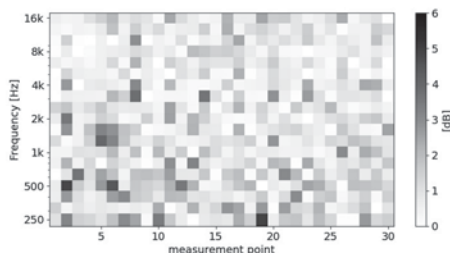


Fig.1: The measurement error between the conventional method and the continuous acoustic measurement method.

1-R-42

1-R-42 インパルス応答推定における直交性に着目した計測用駆動信号設計の一検討
 Driving Signal Design Focusing on Orthogonality in Impulse-Response Measurement
 ☆徳富 響(NBU), 沖田 和久(NBU 卒(現 NJMC)), 近藤 善隆(J-Tec), 福島 学(NBU), 松本 光雄(O), 柳川 博文(arsl)

- ◆目的: 短時間インパルス応答 (IR) 推定値のノイズレベル推定
- ◆これまで: DLR-CS 法で統計的ノイズレベル推定が可能な事を報告
- ◆理論: $y(t) = x_1 * h(t) + x_2 * h(t)$, where $x(t) = x_1(t) + x_2(t)$
 $\tilde{h}(t) = (h(t) + g_1(t)) + (h(t) + g_2(t))$
 $\gamma_{x_1 x_2} = 0$ ならば、 x_1 で推定した h_1 と x_2 で推定した h_2 の差分から「誤差 (ノイズレベル)」が求まる。
- ◆検証: 一様分布ランダム信号 (Random Seq.) (相関係数 0.0001) 位相角 $\pi/2$ の位相差の純音から作成した TRaPS (Fig.1)
- ◆結果: ノイズレベル推定精度は両者でき、TRaPS の方が良い (Fig.2)

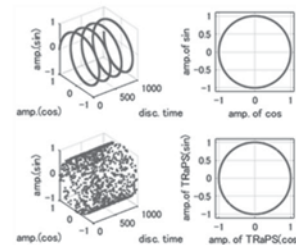


Fig. 1 Lissajous Figures of Two Orthogonal Relationships
 Top: Relationship between sine and cosine

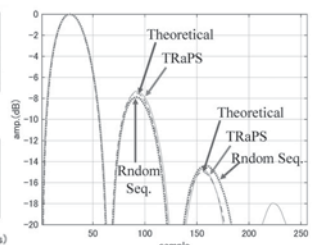


Fig.2 Comparison of Estimated IRs
 Blue dashed line: h (Theoretical)
 Red dotted line: Est. IR (Random Sequence)

1-R-44

1-R-44 高速度カメラと構造化照明を用いたスネアドラム打面の振動の可視化
 Visualization of vibrations on a snare drum surface using a high-speed camera and structured light
 ☆安藤元暉, 新美智也, 山田宏樹, 矢田部浩平(農工大)

- 手法** 構造化照明を投影し高速度カメラで撮影
 ▶ 膜の振動に影響を与えず振動を解析
- 調査** カラー / モノクロ, 単方向 / 複数方向の縞を投影
 ▶ 構造化照明の違いによる情報の変化を調査
- 結果** 単方向の縞を用いて空間的に解析することが可視化に優れた手法であることを確認した

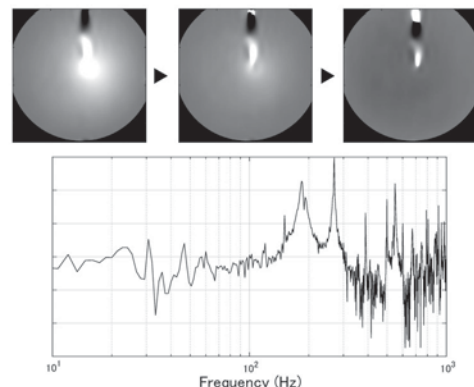


Fig. 1 Visualization of snare drum vibrations and spectrum by Fourier Transform Profilometry

1-R-47

1-R-47 開放型ヘッドホン及びオープンイヤー型イヤホンにおける耳との音響接合と3次元音響再生精度の関係

Relationship between acoustic coupling to the ear and 3D sound reproduction accuracy in open headphones and open-ear earphones.

◎千葉大将, 加古達也, 渡邊悠希, 野口賢一 (NTT), 中村風香, 飯田一博 (千葉工大)

- ◆バイノーラル録音・再生において, 3次元立体音響再生精度の劣化を防ぐための補正フィルタでは, ヘッドホンやイヤホンの装着による音響インピーダンスの変化を表す, PDRと呼ばれる項が含まれる.
- ◆PDRは個人差や装着位置の影響があるため, PDR項の補正を省略できるようにPDR≒1である機種が望ましい.
- ◆耳を塞がないオープンイヤー型のヘッドホンやイヤホンはPDR≒1の可能性が高いが未調査である. そこで, 耳との音響接合が異なるオープンイヤー型の機種種のPDRを測定・評価した.
- ◆100 - 10 kHz では, 耳甲介腔や外耳道入口に筐体が深く入り込まないオープンイヤー型の機種でPDRが1に近い傾向を確認した.

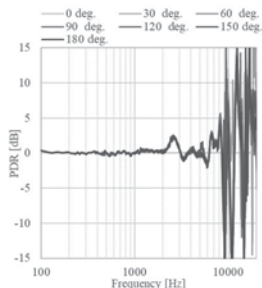


Fig.1: PDRs for each direction for nwm ONE.

Devices ϕ (deg.)	0	30	60	90	120	150	180	Mean	RMS
nwm MWE001	1.8	0.8	1.9	1.1	0.9	0.9	1.2	1.2	5.0
Sony STH40D	4.2	3.4	4.2	3.7	3.4	3.5	3.2	3.6	4.0
Shokz OpenFit	2.2	1.9	2.2	1.9	2.0	1.9	2.0	2.0	3.0
Cleer ARC2	1.7	1.7	1.6	1.7	1.9	1.6	2.2	1.8	2.0
Bose Earbuds	1.2	1.6	1.2	1.3	1.7	1.4	2.0	1.5	1.0
ambie AM-01	0.8	0.6	1.0	0.4	0.8	0.4	1.2	0.7	0.5
ATH-CM707	5.3	4.9	5.3	5.1	4.9	4.9	4.8	5.0	4.0
ATH-EW9	2.5	1.9	2.5	2.1	2.0	1.8	1.9	2.1	3.0
nwm ONE	0.6	1.2	0.6	0.9	1.4	0.9	1.7	1.1	1.0
nwm ONE (offset)	1.5	2.4	1.5	2.1	2.5	2.3	2.9	2.2	2.0
DT990 pro	1.6	0.8	1.7	1.0	1.0	0.7	1.0	1.1	1.0
AKG K712	3.0	1.8	3.1	2.2	1.7	2.0	1.4	2.2	2.0
Sony MDR-MV1	2.2	1.4	2.2	1.7	1.6	1.4	1.5	1.7	1.5
Sony MDR-M1ST	2.2	1.4	2.2	1.6	1.4	1.4	1.4	1.7	1.5

Fig. 2: RMS value of PDR for each device.

1-R-49

講演取消

1-R-48

1-R-48 立体音響再現のための耳介形状に依存しないマルチチャンネルヘッドホンの再生・測定系の検討

Investigation of playback and recording systems for multi-channel headphones that do not depend on the shape of the pinna for reproducing spatial sound.

○加古達也, 千葉大将, 小塚詩穂里, 伊藤弘章, 野口賢一, 阪内澄宇(NTT) 羽田陽一 (電通大)

- ◆ 耳介を覆うマルチチャンネルヘッドフォンを新たに作成し, 複雑な本人の耳介形状を直接利用した立体音響再現を提案
 - HRTF は耳介形状の違いにより個人性が大きい
 - 耳近傍にマイクアレイを配置し, 耳介の影響を除いた頭胸部形状伝達関数(HCSTF)を観測し, マルチチャンネルヘッドフォンで観測したHCSTFを再生することで立体音響再現を行う
- ◆ HRTF を利用しなくても汎用HRTFと同程度の定位性能
 - 22名の被験者で音像定位の主観評価試験を実施し, ダミーヘッドで測定した汎用HRTFと同等の方位角, 仰角の定位性能を確認



Fig.1 Wearing multi-channel headphones with 18 speakers.

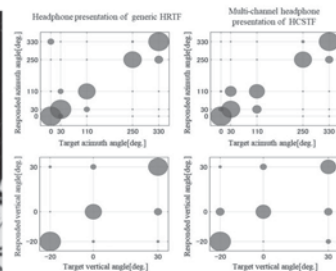


Fig.2 Sound localization using multi-channel headphones that do not use HRTF.

1-R-50

1-R-50 直線スピーカアレイを用いたインタラクティブ焦点音源の自己所有感と定位との関係について

Relationship between sense of ownership and localization of interactive focused sound source using a linear loudspeaker array

☆末藤立己, 澤尻晃大, 羽田陽一 (電通大)

- ◆ 円を描くように指示した手に合わせて移動する焦点音源に対し, 音像移動パターンごとの定位感, 操作感, 一体感を評価した.
- ◆ 主観評価実験には, 32ch直線スピーカアレイとモーションキャプチャを用いた.
- ◆ 音像移動パターンは手に追従 (Follow), 手に遅れて追従 (Delay), 手から π だけ回転 (Rot- π), 手から $\frac{\pi}{2}$ だけ回転 (Rot- $\frac{\pi}{2}$), ランダムに回転 (Random) の5つとした.
- ◆ Fig. 1 は音像位置と手の位置の一致度 (Location matching), 音像と手の一体感 (Sense of unity) についての平均評価値の結果である.
- ◆ 結果より, 音像と手の一致度が低いと, 音像と手の一体感が低下する傾向があることがわかった.

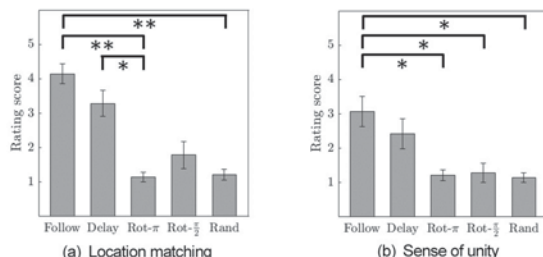


Fig.1: Mean rating score

*p<0.05, **p<0.01

1-R-51

1-R-51 2つのインタラクティブ仮想音源の
 コヒーレンス制御による分離・結合した音像の呈示
 Representation of Separated and Combined Sound Images through
 Coherence Control of Two Interactive Virtual Sound Sources

☆澤尻晃大, 羽田陽一(電通大)

- ◆ 42ch 球面スピーカアレイによるリアルタイムNFC-HOAを用いて、両手の位置に仮想音源を呈示するシステムを構築し、2仮想音源間のコヒーレンスを制御することによって音像の分離・結合を制御可能かを検討した。
- ◆ 水平・垂直・奥行方向に両手を対称的に動かしながら、種々のコヒーレント度を持つ2仮想音源を受聴してもらい、定位の様子が変わったと感じる位置を逐次記録した。
- ◆ 一例として、コヒーレント度 $\kappa = 0.65$ において、手を水平に広げた際の開き角に対する定位の回答確率を Fig. 2 に示す。開き角が大きくなると、1音源として定位 (Joint) していたものが 60° を境に分離 (Sept.) することがわかった。また、一定確率で1つの音源として上方に定位 (Upper) する傾向も見られた。



Fig. 1 Virtual sound sources manipulation by opening arms horizontally

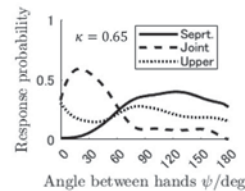


Fig. 2 The response probabilities for each "localization type" when extending both hands horizontally