

### 1-1-1

#### 1-1-1 独立深層学習テンソル分析に基づく多チャンネル音源分離

Determined audio source separation using independent deeply learned tensor analysis.

☆成澤直輝(東大), 池下林太郎(NTT), 高宗典玄(東大),

北村大地(香川高専), 中村友彦, 猿渡洋(東大), 中谷智広(NTT)

- ◆独立深層学習行列分析 [Makishima+, 2019] では, 音源モデルとして周波数間相関を考慮していない.
- ◆本研究では周波数共分散を DNN で推論する多チャンネル音源分離手法である独立深層学習テンソル分析 (IDLTA) を提案する. 周波数間相関をモデル化することでより詳細なモデル化が可能となり, 分離性能の向上が期待できる.
- ◆周波数共分散を DNN で直接推論するのは困難なため, パワー推定 DNN の出力である対角行列と波形推定 DNN の出力であるランク 1 行列の凸結合による周波数共分散のモデル化を提案する.
- ◆音源分離実験によって提案手法の優位性を示す.

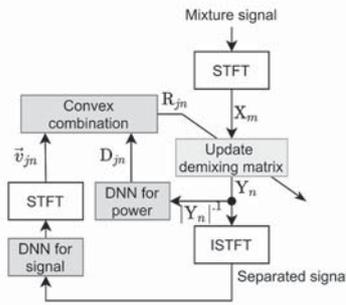


Fig. 1: Separation principle of IDLTA.

### 1-1-3

#### 1-1-3 潜在アナログフィルタ表現に基づく畳み込み層を用いたサンプリング周波数非依存な DNN 音源分離

Sampling-frequency-independent DNN audio source separation by convolution layer with latent analog filter representation

☆齋藤弘一, 中村友彦(東大), 矢田部浩平(早稲田大),

小泉悠馬(NTT), 猿渡洋(東大)

- ◆音源分離は様々なアプリケーションの前処理として利用されることが多く, 様々な音響信号でも動作することが望ましい.
- ◆サンプリング周波数は音響信号の重要な構成要素の1つであり, アプリケーションに応じて決まる.
- ◆しかし, 従来のdeep neural network (DNN) 音源分離手法では学習データによって暗にサンプリング周波数が仮定されており, 学習に用いたサンプリング周波数以外の音響信号には適用できなかった.
- ◆本稿では, 任意のサンプリング周波数の音響信号を扱えるサンプリング周波数非依存畳み込み層と, 提案層を用いたDNN音源分離手法を提案する.

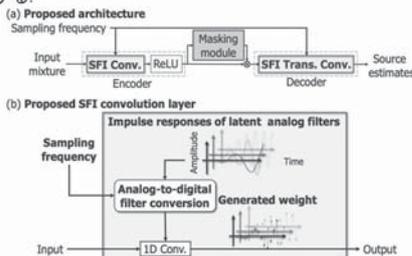


Fig. 1: Architectures of (a) proposed model, and (b) weight generation process of proposed sampling-frequency-independent (SFI) convolution layer.

### 1-1-2

#### 1-1-2 スペクトログラム無矛盾性をを用いた独立低ランク行列分析の実験的評価

Experimental evaluation of consistent independent low-rank matrix analysis

○北村大地(香川高専), 矢田部浩平(早稲田大)

優決定ブライント音源分離の「スペクトログラム無矛盾性に基づく独立低ランク行列分析 (consistent ILRMA)」を実験的に調査する. 特に, 分離信号のスペクトログラム無矛盾性を担保する更新の直前に, プロジェクションバック (分離信号のスケール任意性を解消する処理) を適用する場合としない場合の性能を比較し考察を加える.

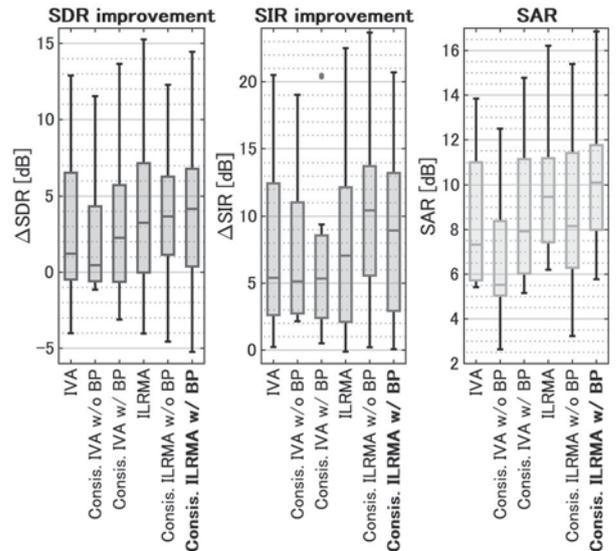


Fig. 1: Evaluation scores for each method.

### 1-1-4

#### 1-1-4 ARMA-FastMNMF に基づく同時的ブライント音源分離・残響除去

Joint Blind Source Separation and Dereverberation Based on ARMA-FastMNMF

◎関口航平(理研 AIP/京大), 坂東宜昭(産総研/理研 AIP), Aditya Arie Nugraha, Mathieu Fontaine(理研 AIP), 吉井和佳(京大/理研 AIP)

- ◆ARMA 過程に基づく残響モデル生成モデルに基づく音源分離手法 FastMNMF に統合し音源分離と残響除去を一挙に行う手法を提案
- ◆NMF に基づく音源モデルに従う各音源信号に対し, 移動平均 (MA) 過程に従って初期残響 (音源に依存) が付加され, それらが重畳した混合信号に対し, AR 過程に従って後部残響 (音源に非依存) が付加される生成モデルを仮定

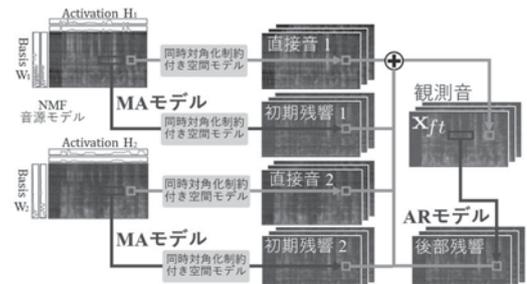


Fig. 1: The generative model of the reverberant mixture signals

- ◆直接音と MA 過程に基づく初期残響の空間相関行列を同時対角化可能な行列に制限することで, 同時対角化行列と AR フィルタ係数を効率的に同時最適化し, 計算量を大幅に削減

### 1-1-5

#### NF-IVAに基づく 線形時変型決定系ブラインド音源分離

Determined Blind Source Separation Based on NF-IVA  
with Time-Varying Linear Transformations

©Aditya Arie Nugraha (理研 AIP), 関口航平(理研 AIP/京大), Mathieu Fontaine (理研 AIP), 坂東宜昭(産総研/理研 AIP), 吉井和佳(京大/理研 AIP)

- ◆ 従来法: 独立成分分析 (ICA), 独立ベクトル分析 (IVA)
  - 線形時不変型決定系 BSS 手法
- ◆ 提案法: 正規化フローに基づく独立ベクトル分析 (NF-IVA)
  - 線形時変型決定系 BSS 手法

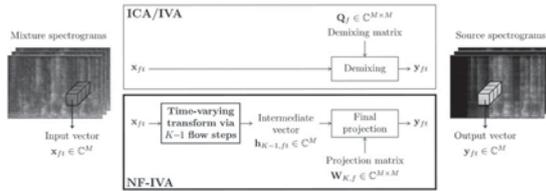


図1: ICA/IVAおよびNF-IVAにおける分離行列の構成方法

- ◆ 体積保存制約を課すことによりパラメータ推定を安定化

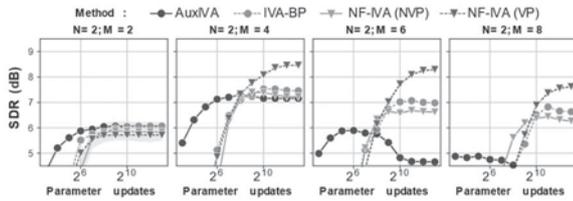


図2: 音源数N・マイク数Mにおける分離音の平均SDR [dB]

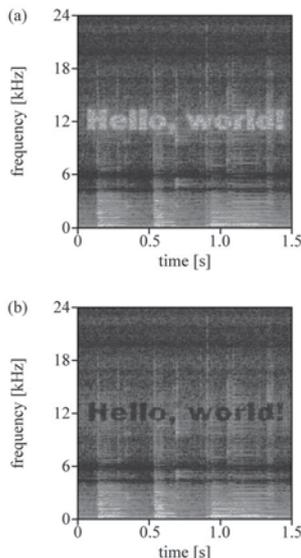
### 1-1-7

#### 1-1-7 スペクトログラムアート通信における テキストメッセージの視認性に関する一評価

Evaluation on the Visibility of Text Messages  
in Spectrogram Art Communications

☆穴沢 早紀, 青木 直史, 尾関 剛成, 土橋 宜典(北大)

本研究では、音波通信の可能性をさらに広げることを目的として、スペクトログラムアートを利用したビジュアルコミュニケーションの可能性について検討している。



スペクトログラムアート通信: (a) 加算方式, (b) 減算方式

### 1-1-6

#### 1-1-6 ランク制約付き空間共分散行列推定法 における補助関数法に基づく雑音 欠落ランク空間基底に対する新しい更新則

New update rule of deficient basis of noise spatial covariance matrix  
based on auxiliary function technique  
for rank-constrained spatial covariance matrix estimation method

☆近藤祐斗, 久保優騎, 高宗典玄(東大),  
北村大地(香川高専), 猿渡洋(東大)

- ◆ ランク制約付き空間共分散行列推定法 (RCSCME) [Kubo+, 2019]は、拡散性雑音中の方向性目的音源をブラインドに抽出するための、前段処理で求めたランクが欠如した雑音空間相関行列の補完に基づいた音源分離手法である。
- ◆ 従来の補完される基底をパラメータとみなしたモデルにおける RCSCME の更新則は EM アルゴリズムに基づいていた。本稿では、同モデルにおいて補助関数法に基づく別の更新則を提案する。
- ◆ 実験により、提案法が従来法より高い分離性能を示すことを確認する。

Table 1: SDR improvements [dB] for each method and noise case. Each term represents "peak score / score after 200 iterations"

Method	babble	station	traffic	cafe
ILRMA	6.1/-	6.2/-	4.7/-	6.4/-
Fixed-basis EM	8.4/7.5	9.5/8.9	7.6/7.2	9.6/8.9
Fixed-basis MM	8.4/7.6	9.7/9.1	7.8/7.5	9.7/9.0
Updated-basis EM	<b>8.9/8.6</b>	10.8/10.6	8.8/8.7	10.6/10.3
Proposed updated-basis MM	<b>8.9/8.6</b>	<b>10.9/10.7</b>	<b>8.9/8.9</b>	<b>10.7/10.4</b>

### 1-1-8

#### 1-1-8 マスクが日本語母音の発音に及ぼす 影響の観測(第二報)\*

E Measurement of the change of the Japanese vowel pronunciation by  
influence of the face mask (the 2<sup>nd</sup> report)

○秋田昌憲, 工藤悠夏, 伊藤智史, 緑川洋一(大分大)

- ◆ 新型コロナウイルスの蔓延によりその着用が強制されているマスクの発音に与える影響を調査している。
- ◆ 先の12月の電気音響研究会における報告では、不織布マスク着用の場合下図のように母音フォルマントのうち第一フォルマントだけが影響を受ける傾向が見られた。ただ、多数話者の測定結果を平均すると傾向がはつきりしなくなる様子も見られた。

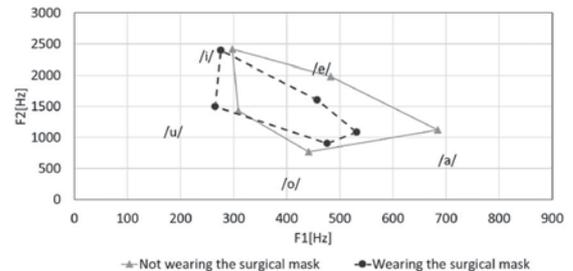


Fig.1 Example illustrating of the distribution of Japanese 5 vowels in case that the examinee wear the surgical mask

- ◆ ここでは、話者ごとにフォルマント移動の傾向を分析し、不織布マスク着用の不快感アンケートとフォルマント移動の結果を整理する。また、不織布マスク以外のマスク着用の追加実験の結果も示す。

### 1-1-9

#### 1-1-9 勾配ブースティング決定木を用いた シャント音識別における特徴量の検出

A Study of Features in Shunt-murmur Identification  
Using Gradient Boosting Decision Trees

☆野田史也, 西島恵介, 古家賢一(大分大)

- ◆腎臓病の患者は日常的にシャントの機能を管理する必要があるが、判定には正しい知識と経験が必要のため、シャント機能を自動的に検査するシステムの実現が望まれている。
- ◆従来の SVM と勾配ブースティング決定木(GBDT)による学習で用いられた特徴量では識別精度が不足している。
- ◆本報告では、特徴量の組み合わせを変更して、5つの学習モデルを作成しアンサンブル学習を行うことで識別精度の向上を図る。
- ◆実験の結果、提案手法を用いることで正解率の向上が確認された。

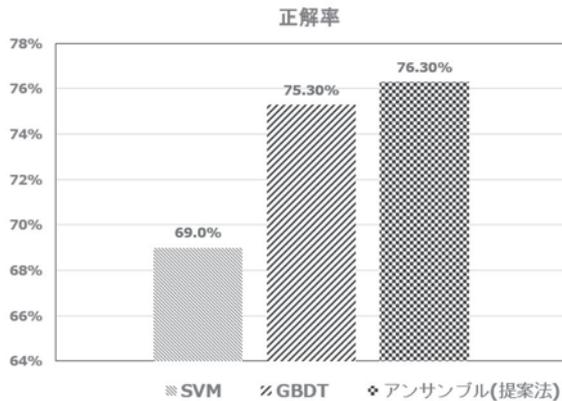


Fig.1: 従来研究と提案手法の正解率の比較

### 1-1-11

#### 1-1-11 稼働音を用いたインジェクタ故障診断

©レ ティエンチエン(日立), 横山 尚, 國岡 昭吾, 猪瀬 聡志(日立建機)

本稿ではスマートフォンで収録した建設機械の稼働音から、AIを用いて複数気筒エンジンの故障有無を診断すると共に、故障がある場合には、詳細診断を通じて故障したインジェクタを特定する方法を提案する。

### 1-1-10

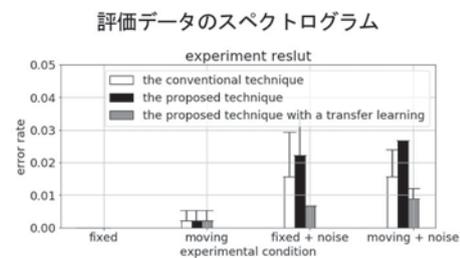
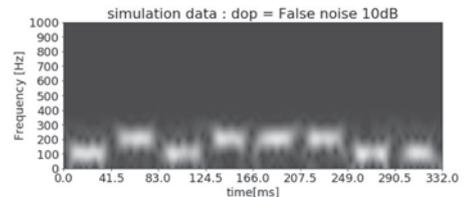
#### 1-1-10 転移学習により設計された 超音波信号識別器の評価

Evaluation of Ultrasonic Signal Classifiers  
Designed by a Transfer Learning Technique

☆尾関 剛成, 青木 直史, 穴沢 早紀, 土橋 宜典(北大)

池田 研一, 安田 寛(株式会社・ソリューション・テクノロジー)

本研究では、非可聴域である超音波信号に情報を埋め込み、データ通信を行うシステムを提案している。本発表では、機械学習を利用して、雑音やドップラー効果といった外乱を含んだ音信設計号から情報を取り出す識別器を設計し、その性能評価を行った結果について述べる。



評価データに対する識別結果

### 1-1-12

#### 1-1-12 Room Acoustic Parameters Estimation using MTF-based CNNs

☆Suradej Duangpummet (JAIST/SIIT), Masashi Unoki (JAIST),

Jessada Karnjana (NECTEC), and Waree Kongprawechnon (SIIT)

- ◆Aim: To propose a blind estimation method for three room-acoustic parameters and an objective index in reverberant environments
- ◆Problem: Inadequate information of room acoustics from an only single estimated parameter or objective index
- ◆Solution: We incorporate the MTF concept into convolutional neural networks (CNNs) to map the temporal amplitude envelope (TAE) of a reverberant speech signal to reverberation time ( $T_{60}$ ) for seven-octave bands ( $T_{60,s}$ ). The approximated MTF and RIR derived from those of the estimated  $T_{60,s}$  are used to calculate the parameters/index (i.e.,  $T_{60}$ ,  $C_{80}$ ,  $D_{50}$ , and STI), as shown in Fig. 1.
- ◆Evaluation: Speech signals in reverberant environments were used.
- ◆Summary: The proposed scheme achieved blind estimation for three room-acoustic parameters/index in reverberant environments.

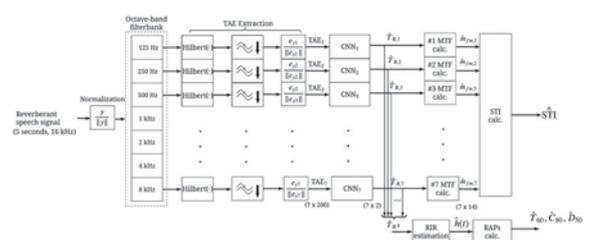


Fig.1: Diagram of MTF-based CNNs method

### 1-1-13

#### 1-1-13 直線型パラメトリックスピーカアレーを用いた復調波の波面合成に基づく焦点音源形成の実環境評価

Evaluation on focused sound construction based on wave field synthesis of demodulated wave using linear parametric loudspeaker array in the real environment

☆佐山史織(立命館大院), 中山雅人(阪産大), 西浦敬信(立命館大)

- ◆我々はこれまで、直線型パラメトリックスピーカ(PAL)アレーを用いた放射音の焦点形成に基づく仮想音源構築手法を提案してきた。しかし、従来の手法では放射音の焦点を形成するため焦点位置ごとに全てのPALの配置角度を調整する必要があった。
- ◆そこで本稿では、PALの配置角度を固定した状態で仮想音源を構築するため、復調波の波面合成に基づく焦点音源形成を行い、性能を評価する。Fig.1に提案手法の概要を示す。焦点音源法により各PALから放射する信号に適用する駆動関数を設計する。そして復調波の波面合成を実現するため、駆動関数を適用した信号を振幅変調し、各PALから同時に放射する。評価実験では、Fig.2に示す実験結果の通り、実環境にて任意の位置に焦点音源を形成可能であることを確認した。

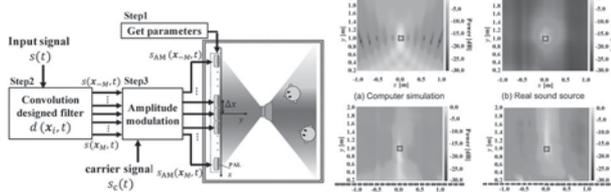


Fig.1: Overview of the proposed method with a linear PAL array

Fig.2: Results for sound pressure

### 1-1-15

#### 1-1-15 耳珠に装着したマイクによる鼓膜位置音場再現について

Reproduction of the sound field at the eardrum position with microphone using tragus.

☆佐瀬 文一(都立大・シス), 大久保 寛(都立大)

3D 立体音響の再現には、頭部伝音関数(Head Related Transfer Function)が有効であることが知られている。HRTF は個人によって異なり、取得法としては、個人の耳型を採取し、作成した耳型マイクによる測定が一般的な方法である。この方法では、耳孔を閉塞した状態で録音を行うこととなる。

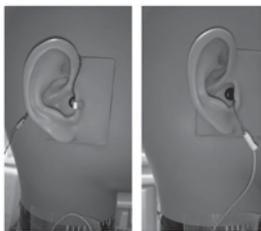


Fig.1 microphone (Left: TC mic, Right: EP mic)

本研究ではHRTF 測定用マイク装置として、耳珠を利用した方法(Tragus-Cuff mic)を提案し、耳孔を開放した状態で測定を試みる。耳珠に装着することによるHRTF への影響と、その方法による鼓膜位置音場の再現性について検討する。

実験の結果、TC マイクは、ヒトの聴こえに影響せず、音の方向の情報を保持した信号を得られることが示された。耳孔を塞がずに音の方向の情報を得られることは、VR のような技術だけでなく、聴覚障害者補聴への応用可能性も考えられる。

また、耳珠を利用できれば多くのデータを簡単に測定できる可能性もあると考えられ、TC マイクの応用と、広く装着可能な方法について検討していく。

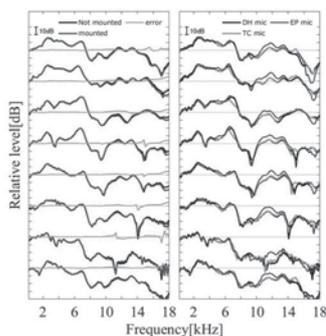


Fig.2 Right: influence of TC mic, Left: Reproduction using each microphone

### 1-1-14

#### 1-1-14 音場合成のための所望音場の事前情報を用いた二乗誤差期待値最小化規準スピーカ配置最適化法

Mean-square-error-based loudspeaker placement using prior information on desired sound field for sound field synthesis

☆木村圭佑, 小山翔一, 植野夏樹, 猿渡洋(東大院・情報理工)

- ◆音場合成のためのスピーカ配置最適化法を提案した。
- ◆重み付きモードマッチング法などの一般的な線形最小二乗法に基づく音場合成法に対して、所望音場の統計的特性を取り入れた二乗誤差期待値規準のコスト関数を定式化した。
- ◆コスト関数を近似的に最小化するための貪欲法に基づく効率的なアルゴリズムを導出した。
- ◆2次元自由空間のシミュレーションにより、提案法で選ばれた配置(Proposed)が経験的に用いられる等間隔配置(Regular)に対して高い再現精度を達成することを確認した。

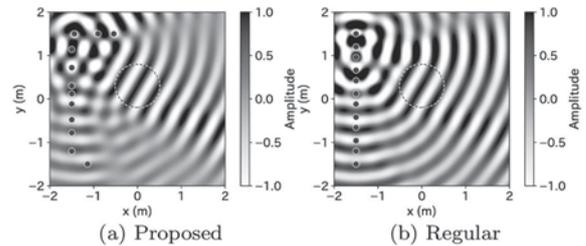


Fig.1:  $-30^\circ$ 方向に伝搬する平面波を重み付きモードマッチング法で再現した際の合成音場と所望音場の音圧分布

### 1-1-16

#### 1-1-16 剛球マイクロホンアレーを用いた境界音場制御の試み

一数值シミュレーションによる再生音場の評価一

Attempt of boundary surface control using a rigid microphone array - Evaluation of reproduced sound field by numerical simulation -

☆福田智也, 伊勢史郎(東京電機大)

- ◆剛球マイクアレーを用いた境界音場制御の原理に基づく音場再現に関する研究事例は未だない。
- ◆本稿では、剛球マイクアレーを用いた境界音場制御の可能性について検討するために、数値シミュレーションによる音場再現精度の評価と再生音場の波面の可視化を行った。
- ◆Fig. 1に中空・剛球マイクアレーを用いた場合の音場再現精度を示す。また、Fig. 2に剛球マイクアレーを用いた場合の波面の可視化を示す。
- ◆Fig. 1より剛球マイクアレーを用いた場合の音場再現精度は、中空マイクアレーを用いた場合と比較して同程度の精度が得られた。Fig. 2より剛球マイクアレーを用いた場合においても、再生音場の波面が原音場の波面を再現できていることがわかった。

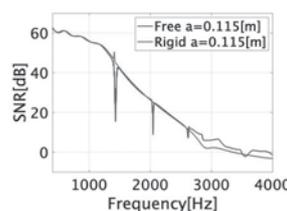


Fig. 1 Sound field reproduction accuracy using free and rigid microphones

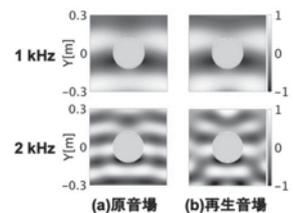


Fig. 2 Sound field with frequencies of 1 kHz and 2 kHz

### 1-1-17

#### 1-1-17 パラメトリックスピーカにおける復調過程の逐次シミュレーションによる復調音の可視化

Visualization of Demodulated Sound Using Sequential Simulation of Self-Demodulation in Parametric Array Loudspeaker

☆下方誠(立命館大院), 中山雅人(阪産大), 西浦敬信(立命館大)

- ◆我々はこれまでに、逐次音線追跡法に基づくパラメトリックスピーカ (Parametric Array Loudspeaker: PAL)の復調音の可視化を検討してきた。しかし、この手法では仮想音源の減衰特性を一般的な点音源の音圧減衰特性と同等にみなしていたため、実測値とシミュレーションした復調音の音圧との誤差が大きかった。
- ◆そこで本稿では、仮想音源の減衰特性を実測した側帯波および復調音の音圧から解析的に算出することで、Fig.1 に示すような実際の仮想音源の減衰特性を考慮した復調音の可視化を実現する。また、本手法では仮想音源の干渉は周波数、波面合成の効果はアレー長に依存することに着目し、周波数、PALの素子数ごとに復調音をシミュレーションする。Fig.2に、評価実験の結果を示す。音圧は収録したTSP信号から算出したインパルス応答をそのまま利用して算出した。Fig.2より、提案手法を用いることで、実環境におけるPALの復調音の伝搬をシミュレーションしたことを確認した。

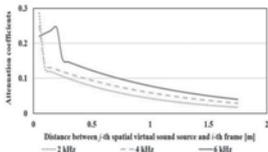


Fig.1: Attenuation coefficients

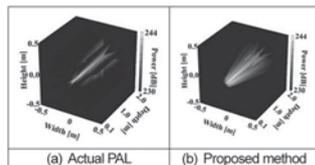


Fig.2: Sound pressure distributions

### 1-1-19

#### 1-1-19 簡易 WFS と VBAP の融合による音源の遠近表現を目指したオブジェクトベース音響レンダリング則

Rendering method for object based audio aiming for perspective expression by an algorithm combined simplified WFS and VBAP

©佐々木陽, 久保弘樹, 長谷川知美, 西口敏行, 杉本岳大, 大出訓史, 大久保洋幸(NHK 技研)

- ◆ナレーションなどの音声オブジェクトと音響メタデータを記録・伝送するオブジェクトベース音響 (OBA) のレンダリング則として、国際規格ではVBAP (Vector Based Amplitude Panning) が採用されている。
- ◆中心以外の聴取位置において、実際の音源では音声オブジェクトの遠近に応じて相対的に音像が定位する方向が異なるが、VBAP で再生される音場では遠近によらず同じ方向になる。(Fig. 1)
- ◆WFS (Wave Field Synthesis) などの音場再現法を利用すれば、音声オブジェクトと聴取位置との相対関係に応じた音像提示が可能である。
- ◆WFS の考え方に基づき、音声オブジェクトの距離に応じた信号分配を取り入れることで、聴取位置との相対関係に応じた音像提示が可能な OBA レンダリング則を提案した。
- ◆簡易な評価実験の結果、所望の音像提示の可能性が示された。



Fig. 1. Difference of a sound localization between an actual sound source and a sound image reproduced by VBAP (left: VBAP, right: Actual)

### 1-1-18

#### 1-1-18 パラメトリックスピーカを用いたハイブリッド 22.2 サラウンドにおける仰角パニングに基づく高解像度音像の構築

Sharper sound image construction based on vertical panning in hybrid combination of 22.2 ch audio with parametric loudspeakers

☆島田 波音(立命大院), 岩居 健太(立命大), 中山 雅人(阪産大), 西浦 敬信(立命大)

- ◆パラメトリックスピーカ (PAL) は動電型スピーカ (EDL) と比べて鋭い指向性を有するため、小さな音像も高解像に構築可能である (Fig.1). これまでに、我々は1台のPALで構築した音像に対して、従来の22.2 ch音響を組み合わせ、空間的に配置された複数のEDLからPALでの再生困難な低域成分を再生することで、所望の位置に音像を保持したまま高音質かつ高解像度な音像を構築可能なシステムを実現した。本研究では、複数のPALを従来の22.2 ch音響と同じ配置で導入し、任意の方向に高解像度音像を構築可能なハイブリッド 22.2 サラウンドを提案する (Fig. 2).
- ◆提案システムの有効性を確認するため、本稿では、提案システムにおける複数のPALを用いて仰角方向へのパニングにより高解像度音像を構築する場合の仰角方向への音像定位性能を評価する。

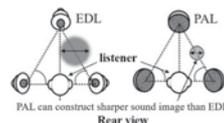


Fig. 1: Overview of sound image constructed with each loudspeaker

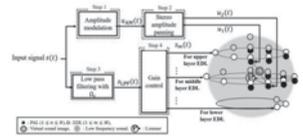


Fig. 2: Proposed system with 12 PALs and 22 EDLs

### 1-1-20

#### 1-1-20 両耳を中心としたマルチゾーン HOA 再生に基づくバイノーラル合成の性能評価

Performance evaluation of binaural synthesis based on multi-zone HOA reproduction at both ears

☆川崎悠季, 大谷真(京大院・工学研)

- ◆HOA 再生では、音場のスイートスポットの半径は周波数に反比例して点に収束するため、頭部中心を再生点とする場合、中心から離れた両耳位置での再現精度は高周波数で低下する。
- ◆本報告では、マルチゾーンHOAに基づき、各耳を再生中心とした二つの再生系によりHOA再生を行う手法を提案する
- ◆剛球モデルにおけるバイノーラル信号について、数値シミュレーションを行った結果、音源が耳に対して反対側にある場合を除き、高周波数での再現精度が向上した。

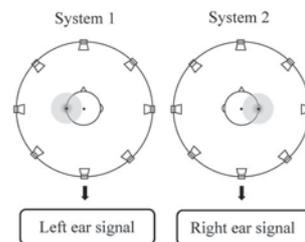


Fig. 1 Proposed method

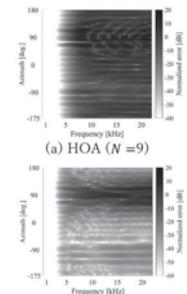


Fig. 2 Normalized Error of Right ear signal

### 1-1-21

#### 1-1-21 畳込みニューラルネットワークを用いた 残響時間が異なる両耳信号による 音源方向推定の検討

Binaural sound localization with convolutional neural network

○西野隆典(名城大)

- ◆立体音響システムの評価指標の一つに音源定位精度がある。ここで、人と同等の評価を行うことが可能なシステムが実現できれば、被験者の負担軽減などに貢献できると考える。
- ◆本研究では、人と同じような音源方向推定が可能となるようなネットワークの構築を目的とし、畳込みニューラルネットワークを用いた音源方向推定を検討した。
- ◆実験では、4層の畳込み層と2層の全結合層から成るネットワークを用い、72方向のクラス分類として音源方向推定を行った。なお、音源信号4種類、残響時間が異なる8種類の両耳室内インパルス応答を用いた。
- ◆実験結果は、 $\pm 5^\circ$ の誤差を許容した場合で約80%以上の正答率となった。ただし、前後誤りが生じないことや、両耳方向での誤差が小さいといった傾向がみられた。

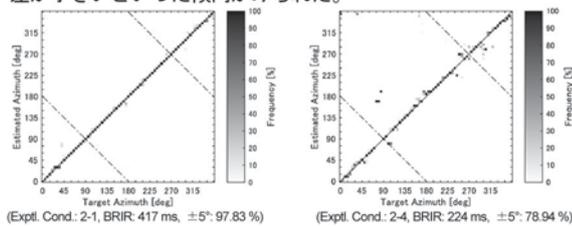


Fig. 1: Results of binaural sound localization.

### 1-1-23

#### 1-1-23 音光変換デバイス「プリンキー」の信号伝搬 過程を考慮した End-to-End 音響シーン分析

End-to-End Acoustic Scene Analysis with Distributed Sound-to-Light Conversion Devices Considering Physical Signal Propagation

◎木下裕磨, 小野順貴(都立大)

- ◆本研究では、我々が研究開発を進めている音光変換デバイス「プリンキー」の音響シーン分析への応用を検討する。
- ◆プリンキーを用いた音響シーン分析においては、音響情報を光により伝送する過程で信号劣化や、カメラに起因する強い帯域制限などにより、十分な音響情報の伝送が難しい課題がある (Fig. 1 参照)。
- ◆そこで、本研究では、プリンキーを用いた音響シーン分析のための最適な音光変換プロセスを End-to-End で学習する、音響シーン分析フレームワークを提案する (Fig. 2 参照)。
- ◆提案するフレームワークは、空気中の光信号の伝播とカメラの応答を微分可能な物理層としてモデル化し、それらを DNN に組み込むことで、最適な音光変換を学習可能とする。
- ◆音響シーン分類のシミュレーション実験により、提案フレームワークは、プリンキーを用いた従来のフレームワークよりも高い分類精度を達成できることが明らかとなった。



Fig. 1: Conventional pipeline for acoustic scene analysis with Blinkies

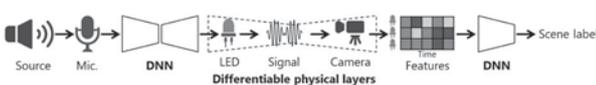


Fig. 2 Proposed end-to-end acoustic scene analysis framework

### 1-1-22

#### 1-1-22 畳込みニューラルネットワークを用いた 昆虫の鳴き声の判別

Insect sound classification with convolutional neural network

△日比野翔, 鈴木千文, ○西野隆典(名城大)

- ◆昆虫の鳴き声は、四季の変化を感じさせる身近な音である。また、同種の個体間で鳴き声によるコミュニケーションが行われており、種類によって鳴き声異なる。そこで本研究では、畳込みニューラルネットワークを用いた昆虫の鳴き声の判別について検討する。
- ◆実験には7種類(セミ科3種, コオロギ科3種, マツムシ科1種)の虫の鳴き声を用いた。鳴き声の波形データを5秒間ごとに区切り、それぞれの昆虫について100個のデータを用意した。
- ◆予備検討として、1)メルスペクトログラムとパワースペクトログラムを入力特徴量として用いた場合の精度の比較、2)データの増し方法の効果の検証を行った。増し方法はタイムシフト、タイムストレッチ、ホワイトノイズ付与、これらの組合せの4種類を用いた。
- ◆予備検討の結果、メルスペクトログラムを特徴量とし、すべての増しデータを用いることが有効であることが確認された。
- ◆増しデータを含めた学習データ1,750、テストデータを350として判別実験を行った結果、判別精度は最大で99.7%(349/350)となり、昆虫の鳴き声の判別が可能であることが確認された。なお、誤判別の理由の一つとして、複数の昆虫の音が混在していたことが考えられる。
- ◆今後の課題として、昆虫の鳴き声が混在しないような録音条件の検討や、対象とする昆虫の種類数を増やした判別実験を行うことが挙げられる。

### 1-1-24

#### 1-1-24 分散マイク・分散カメラの空間位置情報を 活用したマルチモーダルシーン分類

Multi-modal scene classification with distributed microphones, cameras and its position information

◎安田昌弘, 大石康智, 齊藤翔一郎, 小泉悠馬 (NTT)

##### 【シーン理解】

- 音響信号や映像を用いて、人の周囲環境における複数の事象 (=イベント) の連なりからなる場面 (=シーン) を理解する技術
- 特に映像・音響信号を用いたシーンの分類タスクに焦点を当てる

##### 【課題】

- 分散センサを用いた多地点からの観測や、マルチモーダルなアプローチがシーン理解に有効であることが示されてきている
- 一方で、モーダル数や観測地点数が増えるほど入力次元数は増大し、機械学習に基づく手法の適用が難しくなる (次元の呪い)

##### 【提案法】

- 分散センサの空間位置情報を活用することで、各観測地点に特化した低次元な条件付き特徴を抽出し、シーン分類を行う

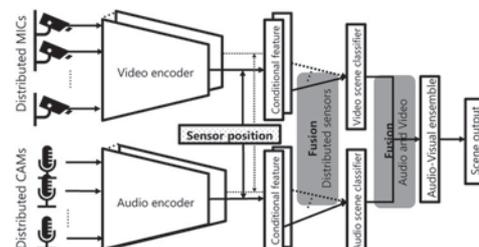


Fig. 1 System overview of proposed method

## 1-1-25

### 1-1-25 Normalizing Flow を用いた自己教師学習による異常音検知

Flow-based self-supervised density estimation for anomalous sound detection

◎土肥宏太, 遠藤隆, Purohit Harsh, 田邊亮, 川口洋平(日立製作所)

- ◆熟練点検者の不足が多くの企業で問題となっており、機械の稼働音を用いて自動で点検を行う自動音響診断システムの需要が高まっている。異常音は取得が困難なため、このシステムでは Variational Auto Encoder(VAE)等を用いて教師なしで異常検知を行うことが多い。
- ◆Normalizing Flow(NF)はVAE と異なり、尤度を正確に推定できるため、教師なし異常検知に有用であるとされる。しかし、滑らかなデータに対して高い尤度を与え、分布外データの検知に失敗してしまう。
- ◆鑑別的自己教師学習は、同じ種類の複数の機械が利用可能な場合に用いられる。稼働音から各機械の機械 ID を判定し、誤判定した場合に異常と判定する。この手法は平均の検知精度を向上させるが、いくつかの機械では検知精度が著しく低下する。
- ◆本研究では、NF を用いた自己教師的な密度推定方法を提案する。提案手法では、検知対象の機械の音(対象音)と、検知対象と同じ種類の機械の音(対象外音)を用い、対象音に高い尤度、対象外音に低い尤度を与えるように学習を行う。この手法は、対象音と同じ種類の機械音を用いる点で自己教師的な手法である。また、対象音の尤度を高くするように学習するため、教師なしの手法に類似している。
- ◆NF として Glow および Masked Autoregressive Flow (MAF) を用いた結果、提案手法は Glow で 5.8%、MAF で 4.6%教師なし手法の検知精度を上回り、鑑別的自己教師学習よりも安定した検知精度を示した。

## 1-2-1

### 1-2-1 VQVAE に基づくリアルタイム波形ベース声質変換の検討

Investigation of real-time generating waveform voice conversion based on VQVAE

☆大西弘太郎, 中庭亘, 松本光春(電通大)

- ◆VQVAE によって高品質なノンパラレル声質変換が出来ることが示されている
- ◆VQVAE をベースとして GAN を用いた音声合成手法を組み合わせることによりリアルタイム変換を目指す
- ◆変換音声についても Discriminator に判別させる Loss、変換前・変換後で音量が変わらない Loss を取ることで変換音声の品質の向上を図る
- ◆主観評価実験において StarGAN-VC2 を超える品質を示す

Reconstruct speech

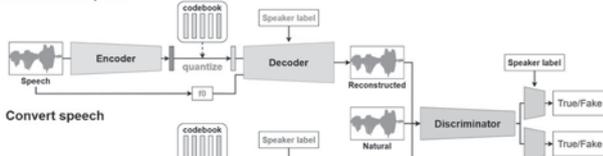


Figure 1 The overview of the proposed method.

Table 1 The MOS score with 95% confidence interval of converted speech.

Model	MOS score
Natural	4.70 ± 0.13
StarGAN-VC2	1.99 ± 0.20
提案手法	3.06 ± 0.27

## 1-1-26

### 1-1-26 Evaluation of Concept Drift Adaptation for Acoustic Scene Classifier Based on Kernel Density Drift Detection and Combine Merge Gaussian Mixture Model

◎ Ibnu Daqiqil Id, Masanobu Abe, Sunao Hara (Okayama University)

- ◆Acoustic scene detection in nonstationary operating conditions make predictive model performance decreases.
- ◆This paper aims to improve and evaluate the CMGMM and KD3 accuracy by tuning the hyperparameter in several concept drift types and scenarios
- ◆Simple and gradual concept drift requires a smaller  $\alpha$  value than recurring because new concept appear continuously, so it needs a high-frequency model adaptation. However, in recurring concepts, the new concept may repeat in the future, so the lower frequency adaptation is better.
- ◆High-frequency model adaptation could lead to an overfitting problem. Implementing CMGMM component pruning mechanism help to control the number of the active component and improve model performance.

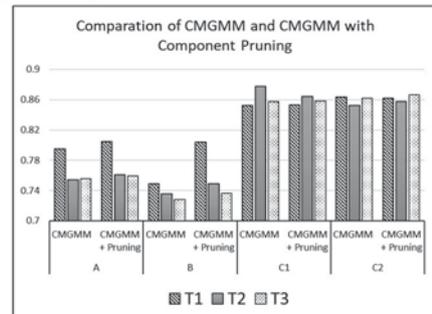


Fig. 1: Performance comparison of CMGMM with component pruning and without component pruning

## 1-2-2

### 1-2-2 MaskCycleGAN-VC: フレーム補間との同時学習による高品質ノンパラレル声質変換

MaskCycleGAN-VC: Improving CycleGAN-VC with frame filling

◎金子卓弘, 亀岡弘和, 田中宏, 北条伸克(NTT)

- ◆メルスペクトログラム領域でのノンパラレル変換の実現を目指し、CycleGAN-VC2 の新しい改良法 **MaskCycleGAN-VC** を提案する。
- ◆本タスクの難しさは、明示的な教師データがない条件下で、メルスペクトログラムに存在する時周波数構造を正確に把握することである。
- ◆解決策として、本研究では、変換の学習と同時にフレーム補間 (**Filling in Frames: FIF**) を学習することを提案する。これにより、自己教師付き学習の枠組みで時周波数構造の学習が可能である。
- ◆主観評価実験では、自然性と話者類似性の二つの評価項目において、**MaskCycleGAN-VC** がベースライン手法の **CycleGAN-VC2** および最新手法の **CycleGAN-VC3** を上回ることを示す。

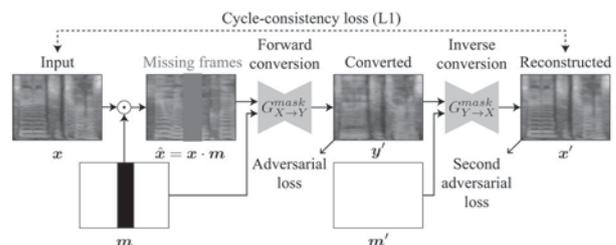


Fig. 1: Overall flow of FIF.

音声サンプル : <http://www.kecl.ntt.co.jp/people/kaneko.takuhiro/projects/maskcyclegan-vc/index.html>

### 1-2-3

#### 1-2-3 CycleVAE 型声質変換を用いた 構音障害者のための高可聴度音声合成 High-Intelligibility speech synthesis for dysarthric speakers with CycleVAE-based voice conversion

◎松原圭亮<sup>(1,2)</sup>, 岡本拓磨<sup>(2)</sup>, 高島遠一<sup>(1)</sup>, 滝口哲也<sup>(1)</sup>, 戸田智基<sup>(3,2)</sup>,  
志賀芳則<sup>(2)</sup>, 河井恒<sup>(2)</sup> (1 神戸大学, 2 情報通信研究機構, 3 名古屋大学)

- ◆ 構音障害者のコミュニケーション支援手法としてテキスト音声合成用いた手法があるが、構音障害者の方からは自分らしい声で話したいという要望が多い。  
一人らしさを持ちつつ可聴性は高い音声を作りたい
- ◆ 構音障害者音声の可聴性改善は難しいが、健常者音声に対して構音障害者の話者性を付与することは比較的簡単と考えられる。
- ◆ 健常者→障害者変換をクロスリンガル声質変換と捉え、CycleVAE 型 VC を導入。健常者 TTS と健常者→構音障害者への変換を行う VC モデルを合わせ、構音障害者の話者性をもちつつ可聴度の高い音声合成可能なニューラル TTS システムを構築。
- ◆ ボコーダには少量データでも学習可能な LPCNet を採用
- ◆ 合成音声を、自然性、話者性および可聴性の3項目で評価

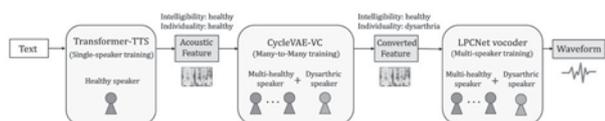


Fig. 1: The flow of the proposed method and training dataset.

### 1-2-5

#### 1-2-5 Pointer-Generator Network に特化した自己教師あり事前学習に基づくテキストスタイル変換 Text style conversion based on masked pointer-generator network for sequence-to-sequence pre-training

◎庵愛, 牧島直輝, 田中智大, 高島瑛彦, 折橋翔太, 増村亮  
(日本電信電話株式会社, NTT メディアインテリジェンス研究所)

- ◆ 本稿では、テキストスタイル変換をタスクとして、Pointer-Generator Network に特化した自己教師あり事前学習方法を提案する。  
➢ 従来の事前学習方法である Masked Sequence-to-Sequence pre-training で用いられている事前学習である token span masking を Pointer-Generator Network の特徴である単語のコピーと生成を効率的に事前学習できるようなタスクに拡張する。

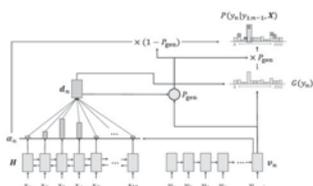


図 1: Pointer-Generator Network

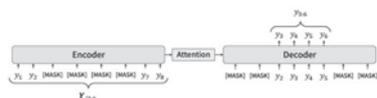


図 2: Token span masking を用いた事前学習タスク

### 1-2-4

#### 1-2-4 CycleVAE を用いた声質変換における WaveNet ボコーダのファインチューニング法の 調査

Investigation on fine-tuning methods of WaveNet vocoder in voice conversion using CycleVAE

☆中谷輝, Patrick Lumban Tobing, 武田一哉, 戸田智基(名大)

- ◆ 声質変換とは、入力された音声の言語情報を保持しつつ、パラ言語・非言語情報のみを変換する技術である。
- ◆ 循環型変分自己符号化器 (Cyclic Variational Autoencoder: CycleVAE) を用いた声質変換では、音声特徴量系列に含まれる言語情報と話者情報を高精度に分離してモデル化することが可能となる。
- ◆ 近年では、変換音声特徴量系列から音声波形への生成過程を従来のボコーダではなく WaveNet ボコーダなどのニューラルボコーダで代替する例が多くみられる。
- ◆ WaveNet ボコーダは学習に大量のデータが必要となるため、様々なファインチューニング手法が提案されているが、学習時と生成時に用いる音声特徴量系列間での質の不一致や、変換音声特徴量系列と目標話者の音声における時間構造の不一致といった問題点が未だ残っている。
- ◆ 先行研究では、CycleVAE の再構成音声特徴量系列を用いることで、高精度な WaveNet ボコーダの学習が可能となり、変換音声の品質を大幅に改善できることが確認されている。
- ◆ 本報告では、CycleVAE を用いたクロスリンガル声質変換における WaveNet ボコーダの学習処理の重要性に着目し、ファインチューニング法に関する詳細な調査を行う。

### 1-2-6

#### 1-2-6 フォトリアル CG エージェントとの マルチモーダル対話システムの構築

Development of a multimodal dialog system with photorealistic CG agent

○北岡教英(豊橋技科大), 西村良太(徳島大), 太田健吾(豊橋技科大),  
原直(岡山大), △石川晃之, △石川友香(TELYUKA),  
△中野涼太(アイシン・ソフトウェア(株)),  
△田中五大, △鍋倉彩那(アイシン精機(株)),  
△佐藤辰耶, △渡邊凌太(アイシン・ソフトウェア(株)),  
△大須賀晋(アイシン精機(株))

- ◆ 将来の人間-機械協奏社会を志向した、リアルな CG エージェントとの音声・マルチモーダル対話システムを構築。  
✓ 3D CG Saya は皮膚や眼球、髪の毛に至るまでハンドメイドで作成され、キャプチャされた実在の人間の動作や表情を基に個性を付与して、独特のゆらぎや不確実性を表現。  
✓ 音声認識・音声合成・赤外線 RGB カメラに基づく画像認識と Saya エージェント出力を、MMDAgent を拡張した対話制御によって駆動。  
✓ 心理学や対話エージェント研究の成果を盛り込み自然なしぐさやインタラクションを実現



Fig.1: Saya as a dialog agent

### 1-2-7

#### 1-2-7 アクセント句を考慮した 日本語 End-to-End 音声合成サーバの構築 Developing a Japanese End-to-End Speech Synthesis Server Considering Accent Phrases

©太田健吾(豊橋技科大), 西村良太(徳島大), 北岡教英(豊橋技科大)

◆日本語の End-to-End 音声合成において自然なアクセントを実現するために、入力文の音素列に加えて、アクセント句の情報も併せて逐次的に入力する手法を提案する。

◆提案手法では、次のように、各音素ラベルの直後に、「アクセント句のアクセント型」と「アクセント核までの距離(音節数)」をそれぞれ表す2種類の整数ラベルを加える(例: “こんにちは”)。

k 5 -4 o 5 -4 N 5 -3 n 5 -2 i 5 -2 ch 5 -1 i 5 -1 w 5 0 a 5 0

◆また、提案手法のサーバ実装についても紹介する。本サーバでは、合成音声の波形情報だけでなく、合成音声の中の各音素の時刻情報も併せて送信されるため、CG アバターやヒューマノイドロボットのリップシンクに用いることも可能である。

◆本サーバは、オープンソースソフトとして一般公開される予定である([https://www.slip.cs.tut.ac.jp/tts\\_server/](https://www.slip.cs.tut.ac.jp/tts_server/))。

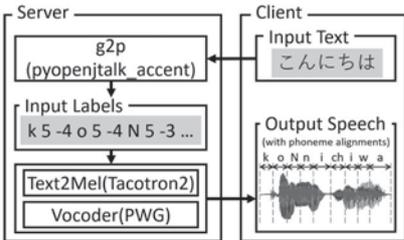


Fig.1: An architecture of our proposed speech synthesis server

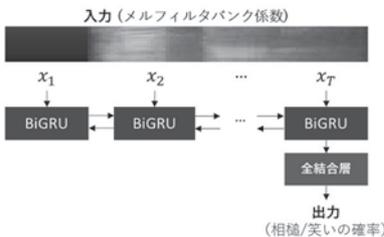
### 1-2-9

#### 1-2-9 ヒューマンロボットインタラクションのための相槌・笑いのリアルタイム検出

Real-time detection of backchannels and laughter for human-robot interaction  
井上昂治, Divesh Lala, 河原達也 (京大・情報学)

##### ■ 相槌・笑いの検出モデル

- 再帰型ニューラルネットワーク (BiGRU) と全結合層
- データ拡張 (SpecAugment) も適用



##### ■ 評価実験

- アンドロイド ERICA の対話コーパス (WOZ)
- F 値は相槌で 90.4%, 笑いで 78.0% を達成
- データ拡張 (SpecAugment) による精度向上

##### ■ 応用

- 相槌と笑いに基づく対話エンゲージメントの推定
- ユーザの笑いに対するシステムの共有笑いの予測

### 1-2-8

#### 1-2-8 ヒューマンロボットインタラクションにおける音響特徴に基づく共有笑いの予測

Prediction of shared laughter based on acoustic features for human-robot interaction  
井上昂治, Divesh Lala, 河原達也 (京大・情報学)

##### ■ ロボットによる“笑い”の生成

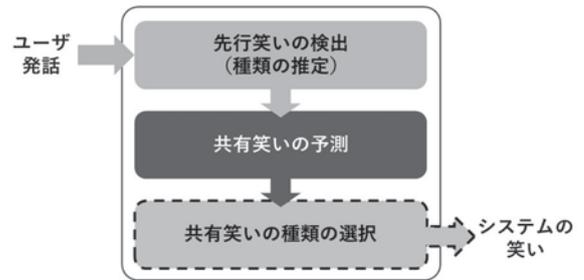
- アンドロイドのような人間らしいシステムには重要
- 適切なタイミング予測には高度な対話理解が必要

##### ■ 「共有笑い」の予測

- ユーザが笑ったときにシステムも笑うか
- ユーザの笑いの音響・韻律・笑いの種類の情報から予測

##### ■ お見合い対話コーパスを用いた評価

- ランダムに予測するよりも高い精度
- リアルタイムで動作した場合の精度も評価



### 1-2-10

#### 1-2-10 TGNN による発話期待度のモデル化に基づく発話タイミング推定

Turn-taking timing prediction based on modeling utterance expectation using TGNN

☆佐久間仁 (早大), 藤江真也(千葉工大/早大), 片山颯人, 小林哲則 (早大)

##### ◆ 概要

- ▶ 多人数会話においてシステムがターンテイキングを行うタイミングを End-to-End に学習可能なモデル TGNN を提案
- ▶ TGNN はユーザの発話末を待たず逐次的に発話期待度を推定し、システム発話のタイミングを決定する
- ▶ 補助タスクとして直前のユーザ発話に対するシステムの応答義務の推定を追加
- ▶ 従来に比べ提案手法がより適切に発話タイミングを推定できていることを確認

##### ◆ 実験

- ▶ 独自に収集した多人数会話コーパスを用いて実験
- ▶ 許容する誤差を変えて precision, recall, f1-score の評価

##### ◆ 結果(Fig.1)

- ▶ 誤差が小さい範囲で提案手法が最も良い性能になった

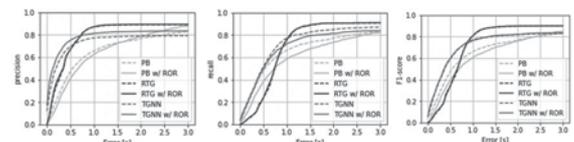


Fig.1: Timing error tolerance - precision(recall, f1) curve

### 1-2-11

#### 1-2-11 言語・F0 特徴量系列を考慮したニューラル音声応答生成の検討

Neural spoken response generation considering linguistic and F0 features.

☆山崎善啓(東北大), 千葉祐弥(NTT CS 研), 能勢隆, 伊藤彰則(東北大)

- ◆ 音声対話システムの多くは応答生成と音声合成のモジュールが独立 → 音声合成時に対話文脈を反映できず単調な韻律で応答してしまう
- ◆ 応答生成時にシステム発話の単語系列と共に **F0 特徴量系列** を生成 → 生成された F0 特徴量を音声合成に用いることで自然な対話韻律の実現が期待される
- ◆ 二つの応答生成モデル (Fig. 1) を提案し, 音声対話コーパスで学習
- ◆ 提案法の連結モデルでは, 自然対話音声により近い F0 系列が得られることが客観的に示唆された。

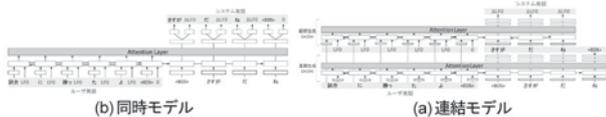


Fig. 1: Architectures of neural spoken response generation model.

Table 1: Objective results.

Model	Encoder input	LF0 RMSE [cent]
同時モデル	L	438.9
	P	442.5
	L+P	433.8
連結モデル	L	407.2
	P	407.8
	L+P	406.2
連結モデル (ユーザ発話 Enc なし)	-	<b>400.0</b>
ピッチ制御なし	-	433.3

L: 言語特徴量, P: 韻律特徴量

### 1-2-13

#### 1-2-13 活性化関数とカーネル関数の関係性を用いたガウス過程音声合成の評価

Evaluation of Gaussian-process-based speech synthesis based on the relationship between activation and kernel functions

○郡山知樹, 猿渡 洋 (東大院・情報理工)

- ▶ 背景: 深層ガウス過程(DGP)に基づく音声合成
  - ・ これまでにDNNより高い性能を報告
  - ・ どこに差異があるか明らかになっていない
- ▶ 手法: 無限幅のベイズNNとDGPの共通構造を利用
  - ・ 様々な活性化関数, モデル構造で比較
- ▶ 実験:
  - ・ 各層で乱択化特徴量に基づくベイズ線形回帰を行う
  - ・ ベイズNN(DBLR)でDGP音声合成と同等の性能を達成
  - ・ DGP音声合成はカーネル関数の選択の影響が小さい傾向

Table 1: Model structure comparison

Model	MCD [dB]	NLL
Non-Bayes DNN	5.11	N/A
MC Dropout	5.13	$2.56 \times 10^2$
BNN	5.08	$2.48 \times 10^2$
RFE-BNN	5.06	$2.48 \times 10^2$
RFE-DBLR	5.01	$2.48 \times 10^2$
DGP	5.01	$2.55 \times 10^2$

### 1-2-12

#### 1-2-12 勾配ブースティング決定木を用いた音声合成手法の検討

An investigation of speech synthesis based on gradient boosting decision trees

☆岩田康平, 高木信二, 橋本佳, 南角吉彦, 徳田恵一(工大)

本稿では, 統計的音声合成の音響モデルに勾配ブースティング決定木を用いた音声合成手法を提案する. 勾配ブースティング決定木は弱学習器に回帰木を用いたアンサンブル手法であり, 勾配を考慮しながら複数の決定木を逐次的に構築することで一つのモデルを構成する. ブースティングを行うことにより単一決定木によるモデルと比べて高精度な予測が可能である. また, 各決定木における計算量は高々木の深さ程度でありニューラルネットワークのように複雑な行列演算を必要としないため, ニューラルネットワークと比較して少ない計算量で特徴量の予測が可能である.

主観評価実験の結果より提案手法においてニューラルネットワークベースの従来手法と同等の音質評価が確認出来た. 今後は時系列方向への考慮を行うことで, より高品質な音声の合成が期待される.

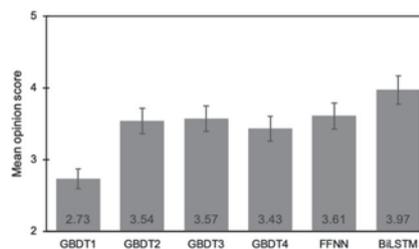


Fig. 1: 主観評価実験

### 1-2-14

#### 1-2-14 HumanACGAN: 人間の知覚を補助分類器に用いた条件付き敵対的生成ネットワークと音素知覚における評価

HumanACGAN: conditional generative adversarial network using human-based auxiliary classifier and its evaluation in representing conditional distribution of phoneme perception

☆上田 陽太(東大院・情報理工), 藤井 一貴(徳山高専),

齋藤 佑樹, 高道 慎之介(東大院・情報理工), 馬場 雪乃(筑波大学),

猿渡 洋(東大院・情報理工)

- ◆ 人間の知覚評価を用いた条件付きの生成モデルを学習する敵対的生成ネットワーク (GAN) を提案する. Fig. 1 に提案法の概要を示す.
- ◆ 従来法の HumanGAN は, GAN の識別器を人間の知覚評価に置き換えた. これにより実在データの分布よりも広い人間が自然だと感じる分布 (知覚分布) を生成器が表現できるようにした.
- ◆ 本研究では HumanGAN の生成器をクラスラベルによる条件付きで学習させる HumanACGAN を提案する. これにより音声合成や音声変換への実用的な応用が可能になる.
- ◆ 生成器はクラスラベルに応じてデータを生成し, 人間は生成データが自然であるか, そのクラスに属するかを評価する.
- ◆ 実験を行い提案法により条件つきで知覚分布を表現する生成器の学習が可能であることを示す.

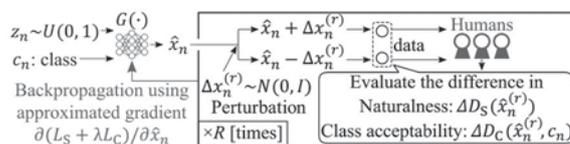


Fig. 1: Training steps of HumanACGAN

## 1-2-15

### 1-2-15 声道のMRI データに基づく DNN 音声合成の予備的な検討

Preliminary study of DNN speech synthesis based on vocal tract MRI data  
 ☆小澤凜夏, 竹本浩典(千葉工大), 平井啓之

- ◆近年, 任意の話者性を再現する音声合成が検討されている。
- ◆そこで, MRI データからフォルマントなどの音声の個人性を表す音響的特徴を抽出して学習することで, 声を失った話者に対しても元の音声の個人性を再現した音声合成ができるのではないかと考えた。
- ◆本研究では MRI データから話者性を再現した音声合成をする前段階の実験として, MRI から抽出できる音響的特徴である母音のフォルマントや, 音声の特徴に寄与する非周期性指標・基本周波数 (F0) などの音響的特徴を埋め込みベクトルとして音声合成し, 話者性が再現できるか検討した。
- ◆主観評価として既存手法の x-vector と提案手法で DMOS 評価をし, x-vector と同程度の話者類似性が得られることを確かめた (Fig. 1)。

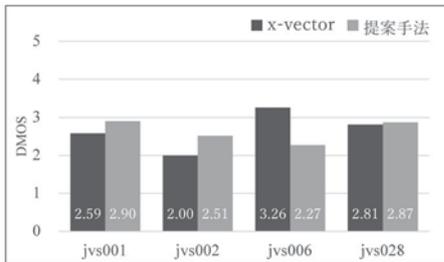


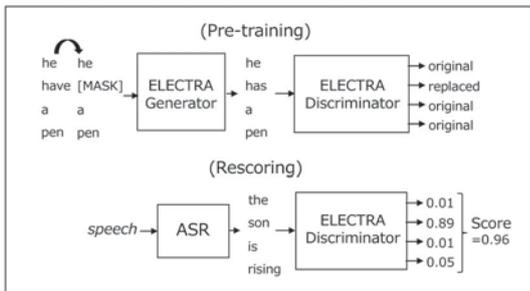
Fig.1: Evaluation result of speaker similarity between x-vector and proposed method

## 1-2-17

### 1-2-17 ELECTRA による 音声認識仮説のリスコアリング

Rescoring Hypotheses of Automatic Speech Recognition with ELECTRA  
 ☆二見颯, 稲熊寛文, 上乃聖, 三村正人, 坂井信輔, 河原達也(京大)

- ◆リスコアリングは従来, 言語モデルによる仮説の尤度に基づく。
- ◆ELECTRA により仮説中の認識誤りを検出し, 誤り数最小の仮説を選択するリスコアリングを提案。
- > 1 回の推論で仮説全体を見ることができる



- ◆Pre-training + 5-best 仮説上の finetuning により, Transformer 言語モデルと比較して一部改善。

Table 1: Rescoring on Librispeech

	WER (%)		runtime
	clean	other	
CTC ASR	6.94	17.12	-
+Transformer LM	4.83	13.93	O(1)
+BERT	4.68	13.73	O(L)
+ELECTRA(finetuned)	<b>4.67</b>	14.06	O(1)

## 1-2-16

### 1-2-16 vq-wav2vec による離散 ID を扱う 音声認識のデータ拡張

Data augmentation for speech recognition using quantized ID by vq-wav2vec  
 ☆上乃聖, 三村正人, 河原達也(京大)

- ◆音声合成を用いた音声認識のデータ拡張は音声合成の学習に多数話者の音声を必要があり, 学習が難しく, また生成時に自然音声と異なる値になるかは保証されない。
- ◆本研究では音声の離散 ID をテキストから合成することで学習の難易度・生成音声と自然音声の差異の問題を解決する。
- ◆音声データの離散化には vq-wav2vec を用い, 音声合成器 (FastSpeech 2) でテキスト系列から ID 列を予測する。
- ◆従来の対数メルフィルタバンク (lmb) を生成するモデルと比較し, vq-wav2vec を生成するモデルは音声認識のデータ拡張の効果が高い。

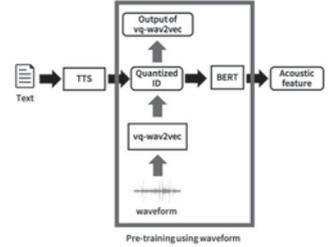


Fig.1: The overall flow of the data augmentation.

Table 1: ASR performance (WER (%)) for Tedium 2. We used 100h paired-data (speech and transcription) from Librispeech100. We generated speech (ID) from Tedium 2 transcription.

Model	dev	test
Real (100h) (lmb) [baseline]	37.95	36.29
Real (100h) (vq-wav2vec) [baseline]	34.39	35.31
Real + TTS (100 + 211) (lmb) [reference]	31.02	32.21
<b>Real + TTS (100 + 211) (vq-wav2vec) [proposed]</b>	<b>25.39</b>	<b>26.25</b>
Real (311h) (lmb) [oracle]	11.70	10.50
Real (311h) (vq-wav2vec) [oracle]	15.42	14.80

## 1-2-18

### 1-2-18 Cross-Modal Transformer に基づく 音声認識誤り訂正の検討

Cross-Modal Transformer based error correction for automatic speech recognition

◎田中 智大, 増村 亮, 庵 愛, 高島 瑛彦,  
 森谷 崇史, 芦原 孝典, 折橋 翔太, 牧島 直輝  
 (日本電信電話株式会社, NTT メディアインテリジェンス研究所)

- ◆音声とテキストを同時に扱う Cross-Modal Transformer に基づく音声認識の誤り訂正手法を提案
- ◆入力音声と音声認識仮説を同時にエンコードすることで両者の情報を考慮したエンコードを実現
- ◆連続値表現に変換した音響特徴量と音声認識仮説の系列を 1 つの系列とみなした後, Self-Attention に入力することにより異なるモーダルの情報を扱う
- ◆自然対話コーパス, および日本語話し言葉コーパスを用いて実験を実施
- ◆音声とテキストを別々にエンコードした場合と比較し, 認識精度が改善することを確認

Table: Character error rate (%)

Model	NTDC	CSJ1	CSJ2	CSJ3
Baseline	13.9	12.3	10.2	11.3
Conventional	13.7	9.8	7.2	7.0
Proposed	<b>13.6</b>	<b>9.4</b>	<b>7.1</b>	<b>6.7</b>

### 1-4-1

#### 1-4-1 ピッチ不安定性強調現象が音知覚に与える影響

Effects of Pitch Instability Enhancement Illusion on Sound Perception

☆川井大輔, 饗庭絵里子(電通大), 津崎実(京都市芸大)

ピアニストの間では、国際標準ピッチで調律されたピアノを演奏した直後に、バロックピッチで調律された楽器を演奏すると、ピッチが揺れるあるいは不安定であるかのような知覚を得ることが知られている (Fig. 1)。本研究においては、この知覚現象をピッチ不安定性強調現象と称し、まずは本現象の存在を確認することと、どのような要因によって生じるのかを検証することで、ピッチ知覚のメカニズムの一端を明らかにすることとした。

そこで、2音から成る西洋音階的である音程と非西洋音階的である音程について、一方の音を5回呈示した後、もう一方の音を3回呈示することで、ピッチが揺れるあるいは不安定であるかのような知覚が生じるかどうか検証を行った。

その結果、音程が西洋音階的であるときは揺らぎや不安定性が減少し、非西洋音階的である場合は増すという傾向がみられ、人間が音程の予測をして音のピッチを知覚している可能性が示唆された。

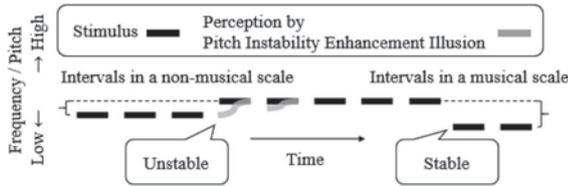


Fig. 1: Pitch Instability Enhancement Illusion

### 1-4-3

#### 1-4-3 周囲環境からの聴覚刺激に伴う個人の集中力の変化に関する初期検討

An investigation into optimizing auditory stimulation with a focus on individuality in order to maintain concentration

○伊藤 弘章, 千葉 大将, 加古 達也, 野口 賢一, 中川 朗(NTT)

- ◆周囲環境音の最適化により、知的作業に集中し生産性を向上するシステム：知覚制御システム（下図）を提案する。
- ◆本システムの実現に向けた以下2点の初期検討結果を示す。
  - 音源種別によって、集中状態及び知的作業のパフォーマンスに及ぼす影響に差異が生じるか？
  - 集中状態及び知的作業のパフォーマンスに影響を及ぼす音源種別に、個人性が影響するか？
- ◆脳波信号からの脳タイプ判定結果を個人性の把握に利用し、聴覚刺激と個人性に基づく集中力の変化について、簡易認知タスク実験により調査した。

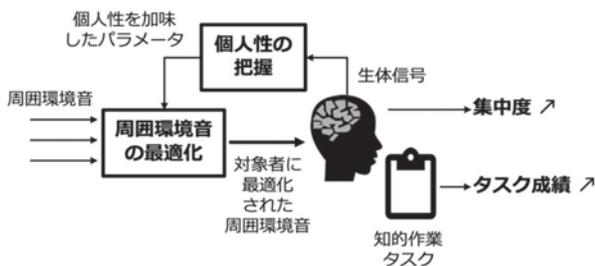


Fig.1: 知覚制御システムの概要

### 1-4-2

#### 1-4-2 サウンドテクスチャ統計量を用いた日常音・感情値の推定

Estimating emotional values of everyday sounds using sound texture statistics

◎中谷裕紀, Δ渡邊雅之, 萬葉穂子(マツダ)

- ◆音で生じる感情の強さ(感情値)の推定を目的として、変調周波数のパワーや周波数間相関(テクスチャ統計量 (Fig.1))により感情値推定モデルを構築した。
- ◆生活音や会話等の日常音データベース (IADS-E) で学習したモデルでは、感情値を推定すると高い相関(相関係数  $r = 0.7, p < 0.01$ )を得た。
- ◆自動車音に適用すると同係数は0.2 ( $p < 0.05$ )と低下するが、自動車音のみで学習すると相関係数は改善した ( $r = 0.9, p < 0.01$ )。

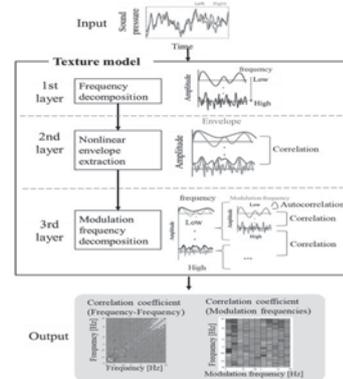


Figure.1 Sound texture model.

### 1-4-4

#### 音韻修復における音声と雑音の角度依存性

The effect of angular difference between intermittent speech sound and inserted noise on phonemic restoration.

☆及川隼平, 坂本修一(東北大学電気通信研究所/大学院情報科学研究科)

- ◆音韻修復の生起要因の一つである連続聴効果には、マスキングが関連することが知られている。
- ◆方向性マスキング解除の先行研究から、断続音声とノイズの位置が離れることでマスキング量が減少し、その結果音韻修復の効果が変化することが予想される。
- ◆そこで本研究ではスピーカアレイを用いて刺激を提示し、音韻修復における音声と雑音の提示角度依存性を検討した。
- ◆実験の結果、高親密度 (7.0-5.5) の単語において、音声の提示方向とノイズの提示方向が  $30^\circ$  以上離れると了解度が減少した。この結果は、方向性マスキングの結果と整合するものである。

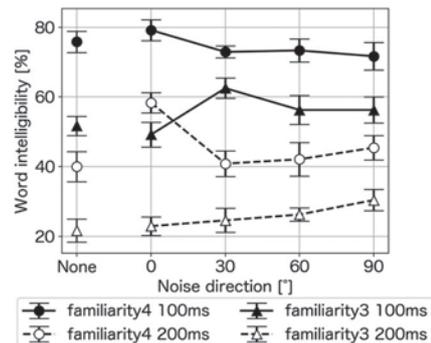


Fig.1: Word intelligibility of intermittent speech sound as a function of word familiarity and angular difference between speech sound and noise

## 1-4-5

## 講演取消

## 1-4-6

## 1-4-6 聴覚的顕著性に寄与する Spectro-Temporal Modulation 情報の検討

◎木所晃利, 木谷俊介, 鶴木祐史(北陸先端大)

- ◆背景: 耳を引き付けられる音(聴覚的顕著性)がどのようなものかは明らかになっていない。
- ◆目的: スペクトル・時間変調(spectro-temporal modulation (STM))の時間変化を手がかりに聴覚的顕著性が説明可能かを検討すること。
- ◆手法: スペクトル情報および時間情報における変調成分を分析可能なSTM分析を用いた。分析の対象は、顕著性スケール(Liao et al., 2016)が求められている10個の刺激を用いた。
- ◆結果: STMの時間変化に着目し分析を行った。STMの時間変化としてコヒーレンスを求め、顕著性スケールとの相関を求めた。その結果、音圧レベルが一定の顕著性スケールとは相関がみられなかったが、ラウドネスレベルが一定の顕著性スケールとの間には有意な相関がみられた(相関係数0.67,  $p < 0.05$ )。

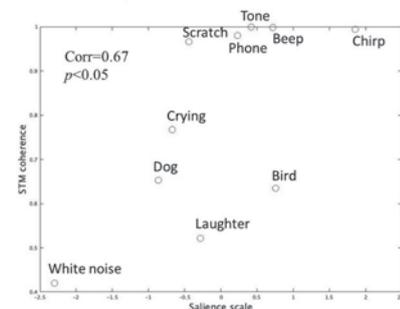


Fig.1: Scatter figure of the salience scale and coherence of the STM

## 1-4-7

1-4-7 げっ歯類モデルにおける  
ミスマッチネガティビティの多点計測

Microelectrode mapping of mismatch negativity in rodents

○白松(磯口)知世, △高橋宏知(東大・情報理工)

生物が聴覚に頼って周囲の環境を探るとき、音の変化に注目することが重要となる。例えば、定常的な環境音に突如紛れ込んだ逸脱音は、鳴き声や足音といった、外敵由来の音である可能性が高い。こうしたことから、非注意下でも音の変化を自動的に抽出することは、生物の生存に有利に働くと考えられる。

我々の聴覚系には、こうした受動的な音変化の抽出機能が備わっている。特に、1978年にヒト脳波で同定されたミスマッチネガティビティ(mismatch negativity; MMN)は、被験者の注意の有無によらず発生するため、脳の自動的な逸脱検出機能に関わる代表的な神経活動として注目されてきた。

MMNの発生機序や聴覚における機能を明らかにするため、ヒト脳波や脳磁図の研究と並行して、げっ歯類を中心とした動物モデルを用いた電気生理学的なMMN計測が盛んに試みられてきた。本発表は、主にげっ歯類モデルで得られた知見として、第一に、MMNの逸脱検出性、すなわち、MMNは単なる刺激への順応では説明できず、繰り返し音からの音変化という文脈に依存して発生していることを述べる。第二に、聴覚におけるMMNの役割の一例として、MMNが個体の経験に依存した音の顕著性を表現している可能性を紹介する。

## 1-4-8

瞳孔・眼球運動反応からの  
注意へのアプローチ

Approaches to auditory attention through eye-metrics

○古川茂人(NTT CS 研)

- ◆感覚・知覚過程で注意は決定的に重要な役割を果たす。しかし、その定量化、メカニズムの解明は未だ十分ではない。
- ◆時々刻々と変化する注意状態を外部から観測するための客観的な指標が求められている。
- ◆瞳孔・眼球運動反応(eye metrics)が注意の様々な側面と関連する神経活動を反映する。
- ◆本講演では、聴覚における外発的注意、内発的注意とeye metricsに関する当研究室の研究事例を紹介する。

### 1-4-9

#### 1-4-9 音声言語認知への単語親密度の影響

Word familiarity effects on spoken language recognition

○天野成昭(愛知淑徳大学)

単語親密度とは、単語の主観的な馴染みの程度を表す指標である。日本語の語彙特性第1巻(天野・近藤, 1999)には、文字呈示、音声呈示、及び文字と音声の同時呈示による約8万語の単語親密度が収録されている。単語親密度は音声言語認知に強い影響を及ぼす。単語親密度が高いほど単語の認知が速く(Fig. 1)かつ正確に(Fig. 2)行われ、音圧低下(Fig. 3)や雑音重畳(Fig. 4)に対する頑健性も高い。また単語親密度は場所や時間に依存しない再現性のある安定した指標である(Amano et al., 2007)。

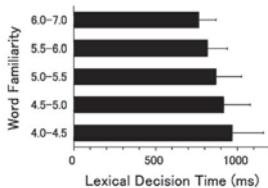


Fig. 1. Lexical decision time as a function of word familiarity.

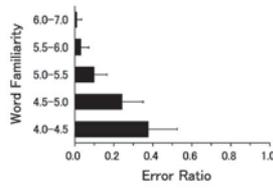


Fig. 2. Error ratio in lexical decision task as a function of word familiarity.

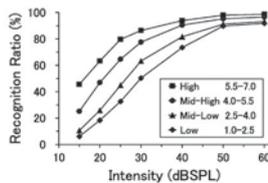


Fig. 3. Word recognition ratio as a function of intensity and word familiarity.

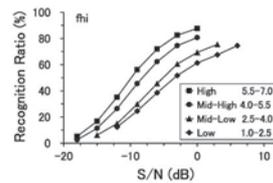


Fig. 4. Word recognition ratio as a function of SN ratio and word familiarity.

### 1-5-2

#### 1-5-2 建築基準法における長屋又は共同住宅の界壁の遮音性能評価に関する改正

Amendment to the Building Standard Law regarding the evaluation of sound insulation performance of separation wall between each unit of row houses or apartment houses

○平光厚雄(国総研)

- ◆令和元年に、長屋又は共同住宅の界壁の遮音性能を規定している建築基準法、建築基準法施行令、同告示が改正された。
- ◆建築基準法第30条第2項では、界壁が小屋裏又は天井裏に達しなくてもよい要件が追加された。
- ◆建築基準法施行令第22条の3では、界壁が小屋裏又は天井裏に達しない場合には、側路伝音音(天井を透過する音)についても界壁と同じ音響透過損失等級R-40相当以上の性能が必要となること規定された。
- ◆昭和45年建設省告示第1827号では、遮音性能を有する天井の構造として、9.5mm厚のせっこうボードに100mm厚以上の吸音材を全面に敷設する仕様が追加された。さらに令和2年の改正により、吸音材を全面に使用しない強化天井の場合は、開口面積に応じた材料を設けることが示された。(Table 1)

Table 1 Effective structure of the opening for sound insulation

開口面積	材料
100cm <sup>2</sup> 未満(開口面積の合計が天井の面積の0.4%以下であるものに限る。)	厚さ50mm以上の吸音材(密度40kg/m <sup>3</sup> 以上のロックウール、密度24kg/m <sup>3</sup> 以上のグラスウール等)またはこれと同等以上の性能を有する材料
上記以外	強化天井と同等以上の遮音性能を有する材料

### 1-5-1

#### 1-5-1 集合住宅における遮音性能の評価方法と住宅性能表示制度に関する考察

Study on Performance System and Sound Insulation Performance of Multi-family Housing

○井上勝夫、阪本一生(日大理工)、平光厚雄(国総研)

住宅品質確保促進法の目的は、欠陥住宅問題の未然防止や工事の請負契約及び売買契約の適正化、住宅の品質に対する意識の向上などを推進することであり、住宅性能表示制度において住宅の性能に関する表示基準、評価の制度を設け、紛争の処理体制を整備することにある。また、新築住宅の請負契約又は売買契約における担保責任について、特別の定めをすることにより、品質確保の促進、購入者利益の保護、紛争の迅速かつ適正な解決を行うこと(第1条)としている。住宅性能表示基準では、住宅の性能に関し表示すべき事項及び表示方法の基準を示し、性能評価基準において、性能表示基準に従い表示すべき性能に関する評価の方法の基準を示している。性能表示基準(国交省告示第1346号、平成15年4月30日)では、音環境に関する表示すべき性能項目として、(1)重量床衝撃音遮断性能、(2)軽量床衝撃音遮断性能、(3)界壁の透過損失等級、(4)外壁開口部の透過損失等級の4項目を掲げている。

本稿は住宅品質確保促進法で示している住宅性能表示制度の音環境に関する項目の普及状況及び今後の方向性について解説するとともに、集合住宅における音環境性能のうち、各部の遮音性能に関する主要項目の評価方法について、今後の方向性について検討したものである。

### 1-5-3

#### 集合住宅の遮音性能評価に関する日本建築学会の規準

Standard of Sound Insulation in Apartment Houses of Architectural Institute of Japan

○田端淳(大成建設)、平松友孝(音環境研究所)

「集合住宅の遮音性能・遮音設計の考え方(日本建築学会編, 2016.1)」の刊行に至る経緯、概要、刊行後の状況について報告する。

##### 【経緯】

集合住宅の普及に伴い、顕在化してきた音環境問題に対応するため、日本建築学会では、関連JIS化に合わせ、1997年に「建築物の遮音性能基準と設計指針(第二版)」を刊行した。同書は長きに亘り遮音性能基準として位置付けられている。その後、集合住宅を取り巻く環境の変化に対応すべく、日本建築学会環境規準として新たな規準の制定を目指して作業を進めたが、規準制定にあたっては、十分な技術的検討・議論と、関連する研究者、技術者、実務者の理解が必要であることから、前段階として「集合住宅の遮音性能・遮音設計の考え方」という形で刊行された。

##### 【概要】

同書で対象とする音環境性能は、室内静ひつ性能、空気音遮断性能、床衝撃音遮断性能である。このうち、以下の2つの性能については、従前とは異なる測定、評価方法を規定している。

室内静ひつ性能は、A特性音圧レベルを測定・評価量としている。

床衝撃音遮断性能のうち、重量床衝撃音遮断性能は、ゴムボール衝撃源を採用し、最大A特性音圧レベルを測定・評価量としている。

##### 【刊行後の状況】

同書刊行後、多くの研究が報告されており、本報告では、変動騒音の評価方法、空気音遮断性能の評価方法、重量床衝撃音遮断性能の評価方法に関する14の研究について、その概要を紹介する。

1-5-4

住宅の音環境性能に関する ISO 規格の最近の動向

Recent development of ISO standards for rating acoustical environment in dwellings

○佐藤 洋(産総研)

- ◆住宅の音環境性能に関する ISO 規格は評価分類規格と部材性能測定・評価規格に大別される。
- ◆評価分類規格としては FDIS 19488 “住宅の音環境性能評価”が国際標準として審議されていたが、2017 年に否決され、2018 年に技術仕様 (TS: Technical Specification) として文書化する方針となり、FDIS を TS として再度文書化するプロセスに入り DTS の段階。もうひとつは ISO DIS 717-2 “床衝撃音遮断性能の格付け”であり、2020 年 10 月に制定された。
- ◆FDIS 19488 に記されていた評価分類は 7 段階であるが、ヨーロッパでは 4 段階の分類がよく用いられている。
- ◆部材性能計測・評価規格は空気音遮断性能の ISO 16283-1、床衝撃音遮断性能の ISO 16283-2、ISO 717-2 および、室内の静ひつ性能の ISO 16283-3、設備騒音の ISO 16032 とその現場計測法の ISO 10052、残響時間計測の ISO 3382-2 がある。
- ◆これらの審議状況は床衝撃音遮断性能の計測・評価にゴムボールを用いることを追加すること除き、従来の規格内容を継続している。
- ◆現状として、住宅に限ったことではないが、床衝撃音遮断性能に関してゴムボールを計測・評価に用いることを規格に盛り込むことが進んでいる。

1-5-5

ビル認証システムにおける住宅を対象とした音環境項目

Acoustic requirements for residential building in building certification systems

○小柳慎一郎(竹中技研)、宮島徹(清水建設技研)

- ◆近年登録件数が増加傾向にある建物の環境性能評価認証システム (ビル認証) を対象とし、集合住宅に関連する音環境関連項目を調査した。
- ◆調査対象: ①WELL v2 ②LEED v4.1 RRESIDENTIAL MULTI FAMILY HOMES、③CASBEE-建築(新築)、⑤CASBEE-住戸ユニット(新築)
- ◆①~④のビル認証で評価される項目の概略を Table1 に示す。室内騒音と戸境の遮音、床衝撃音は今回の調査対象の全てのビル認証で評価される。ただし、LEED での室内騒音の対象範囲は空調騒音のみである。また、外部環境への影響配慮 (音を発生させる側としての評価) は CASBEE のみが評価しており、認証システムによる違いが見られる。Table2 は住戸間の遮音性能基準値の一覧である。③のみ室間音圧レベル差  $Dr$  が基準に用いられることがわかる。
- ◆本講演では評価項目の比較を通して、それぞれのビル認証の特徴を中心に評価項目を解説する。

		①	②	③	④
室内騒音	外來騒音	●	●	●	●
	空調騒音	●	●	●	●
遮音	開口部	●	●	●	●
	戸境	●	●	●	●
	共用部等	●	●	●	●
床衝撃音		●	●	●	●
スピーチプライバシー環境影響の配慮				●	●

Table1: An outline of evaluation items for the building certification systems

認証	評価項目	評価値
①	戸境壁(寝室)	STC55, N1C50
	戸境壁(その他)	STC50, N1C45
②	戸境壁/床/天井	STC50
③	住戸間室	$Dr$ 40未満-55以上
④	住戸間室	(建築基準法の定める透過損失)-Rr-55

Table2: Reference values of sound insulation performance between dwelling units

1-5-6

Partition of Unity FEM による不均質な MPP 吸音体の吸音特性の予測

Predicting absorption characteristics of inhomogeneous MPP sound absorbers with Partition of Unity FEM

☆向江俊一、奥園健、草鹿もどり、田丸加奈子、阪上公博 (神戸大・院)

- ◆Helmholtz 方程式の一般解として平面波を形状関数に導入する Partition of Unity FEM(PW-FEM)を用いて、不均質な MPP 吸音体(Fig. 1)の吸音特性を予測できるかどうかを検証する。
- ◆はじめに、十分な空間分解能を有する空間 4 次精度 FEM を参照解として、PW-FEM による音場予測の妥当性を確認した。
- ◆次に、合成インピーダンス法を理論値とし、実測値と理論値との比較から PW-FEM による吸音特性予測の妥当性を検討する。
- ◆PW-FEM は、理論での予測が難しい不均質性の高い MPP 吸音体についても予測できる可能性を示した(Fig. 2)。

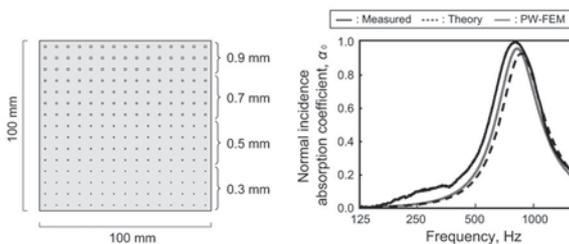


Fig.1 An example of used inhomogeneous MPPs: Specimen B.

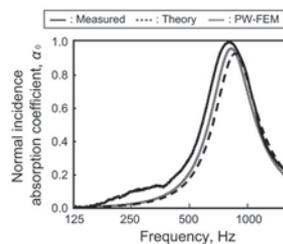


Fig.2 Comparisons among absorption characteristics of Specimen B measured, obtained by the theory and the PW-FEM.

1-5-7

3D プリンターによる造形物の吸音率に関する FE 解析

FE analysis for sound absorption coefficient of object using 3D printer

○黒沢良夫(帝京大)、△中島智、△石橋圭太(東ソニー株式会社)

- ◆自動車や住宅や家電等の騒音低減のため、ウレタンフォームが一般的に用いられている。吸音特性を計算する場合、Biot-Allard モデルが用いられるが、計算で用いられるパラメータが実際の材料開発に結び付きにくい。
- ◆そのため、3D プリンターを用いてウレタンフォームを単純構造化した造形物(テストピース: 図 1)を作成した。数ミリの径の球状の空洞を円柱状の微小な穴でつないだ造形物を球の径や円柱の径が異なる複数種類作成し、音響管を用いて垂直入射吸音率を計測した。また、同じ形状の FE モデルを作成し吸音率を計算した。計測によって得られた知見や計測結果と計算結果の比較(図 2)等を紹介する。

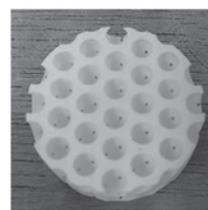


Fig.1: Test piece

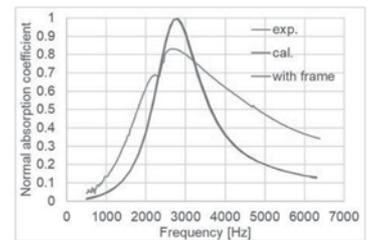


Fig.2: Comparison between experimental result and calculation results

### 1-5-8

#### 1-5-8 単板ガラス窓の遮音性能の実用的予測式の構築 (1)

Study on practical predicting sound insulation performance of a single-leaf window (1)

◎塚本陽平 (YKK AP/神戸大), 玉井芳, 阪上公博, 奥園健 (神戸大・工)

- ◆建築物の音環境を設計するうえで、窓の遮音性能は重要な要素であるが、現在のところ窓の遮音性能の予測技術は十分であるとはいえない。
- ◆本報では既往の板の遮音理論を組み合わせることで窓の遮音性能の予測に適用する方法を検討する。なお、もっとも単純な窓構成である単板ガラスの窓を対象とする。
- ◆Fig.1 に示すようにコインシデンス効果および隙間の影響を受けない低域における Sewell の式と、限界周波数以上の Cremer の式の組み合わせによる予測モデルはFIX窓の測定データと概ね一致する。
- ◆また、Sewell の式を適用する場合には、一般的な範囲の大きさの窓であれば、窓面積  $F=1$  として計算すると測定値とよく一致する。
- ◆引違い窓の場合は中音域～コインシデンス限界周波数の範囲で一致しない。窓の種類による差異への対応が課題である。

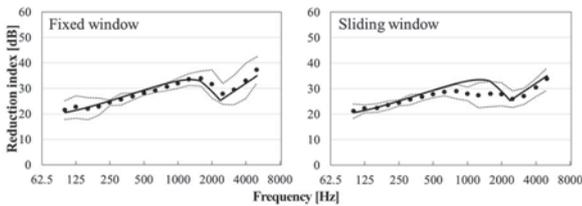


Fig.1: Prediction of reduction index of window with 5 mm glass. (Plots: Average of measurement data, Dotted line: 95% confidence interval, Solid line: Predicting model)

### 1-5-10

#### 1-5-10 低周波音による木造家屋壁面の振動分布推定 -音圧計測と境界要素法に基づく逆問題解析-

Estimation of vibration distribution on wooden house façades by low-frequency sound -Inverse problem analysis with sound pressure measurement and boundary element method-

◎劉金雨 (東大・新領域), 井上尚久 (前橋工大), 佐久間哲哉 (東大・工)

- ◆本報では、実大木造試験家屋において低周波音に対する室内多点で計測した音圧分布から境界要素法逆問題解析により壁面・窓面の振動分布の推定を試み、壁・窓構造の影響を検討した。建築構造は、Case 1: 外壁面密度  $25\text{kg/m}^2$ 、単板5mm一重窓、Case 2: 外壁面密度  $55\text{kg/m}^2$ 、Case 3: 加えて二重窓に交換、3条件を設定した。
- ◆Case 1 では全体的にレベルが高く、周波数の上昇に従って低下の傾きが大きい。Case 2, 3 では部位によるレベル差が大きく、二重窓を設置した Case 3 ではむしろ窓面以外の部位でレベルが低下している。窓の遮音性能の上昇により室内音圧が低下し、壁面振動を抑制したとも考えられるが、現時点では振動分布推定に問題がある可能性も否定できない。

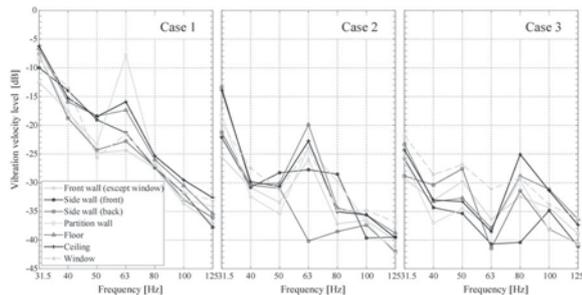


Fig.1: Relative vibration velocity level of each surface of the façades.

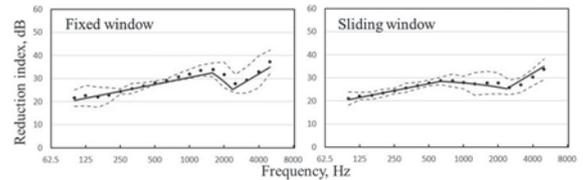
### 1-5-9

#### 1-5-9 単板ガラス窓の遮音性能の実用的予測式の構築 (2)\*

Study on practical predicting sound insulation performance of a single-leaf window (2)

☆玉井芳, 阪上公博, 奥園健 (神戸大・工), 塚本陽平 (YKK AP/神戸大)

- ◆単板ガラスの窓の遮音性能について、前報(1)では理論式に基づく予測法で実測データとの良い一致を示したが、隙間など、窓特有の因子による透過損失への影響の予測に課題を残している。一方、透過損失の傾きを利用した回帰直線の検討より、更に簡単な表現も可能であると考えられる。
- ◆本報では多数の実測データから回帰分析を行った結果を利用し、さらに簡便で、実用的な予測手法を構築することを目指す。窓種の異なる単板ガラス窓を対象とし、回帰直線と Sewell の理論値および Cremer の理論式を用いた予測法を提案し、測定データとの比較からより簡易で妥当性の高い予測手法を見出すことを目的とする。
- ◆FIX窓、たてすべり出し窓については  $100\text{ Hz} \sim 2 f_c/3$  の範囲で、引違い窓については  $100\text{ Hz} \sim 630\text{ Hz}$  の範囲で  $3\text{ dB/oct}$  の直線を用い、 $f_c$  以上では Cremer 式を用いることでいずれも良好な結果が得られた。



Prediction of reduction index of window with 5 mm glass (Plot: Average of measurement, Dotted line: 95% confidence interval, Solid line: Predicting model)

### 1-5-11

#### 1-5-11 排水立て管・継手から発生する流水音の放射特性 -その2 排水高さと負荷流量に着目した実験的検討-

Characteristics of sound radiated from drainage stacks and fittings focused on falling height and flow rate

☆間瀬圭祐, 岡田恭明, 吉久光一 (名城大・理工), 市川友己 (三井住友建設)

- 排水立て管と継手から発生する流水音の測定を実施している。今回、構造の異なる2種類の継手を使用して、排水高さと負荷流量に伴うレベルの変化について検討した。Fig.1 は、実験対象とした中高層(左)と超高層(右)建物用の継手であり、鉛直方向の長さや内部に配置されている減速と旋回流形成のためのガイドの枚数が異なる。両継手を用いて排水負荷流量に対するレベル増加について検討した結果が Fig.2 である。上段が超高層用 ( $1.7 \sim 10\text{ L/s}$ )、下段が中高層用継手 ( $3 \sim 7.5\text{ L/s}$ ) の場合であるが、流水音のレベル増加の傾向は概ね同程度であった。その他、排水高さの位置を  $8.1 \sim 16.2\text{ m}$  まで変化させた時の流水音の周波数特性の比較、立て管あるいは継手を被覆することによる騒音低減効果に着目した検討結果などについても報告する。

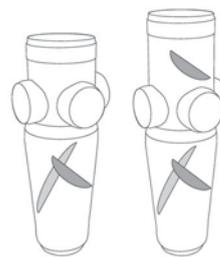


Fig.1 Fittings for medium-high rise and tower buildings.

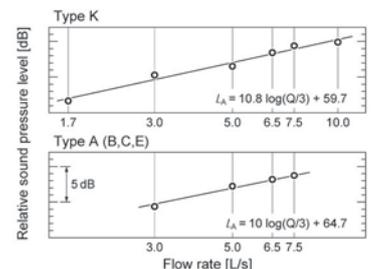


Fig.2 Dependence of flow rate on A-weighted sound pressure level  $L_A$  (PS).

## 1-5-12

### 1-5-12 一次元拡散体の天井吸音を有する会議室の音響指標への影響の時間領域有限要素解析

Time-domain FE analysis of 1D diffusers' effects on acoustic parameters in meeting room having ceiling absorber

○吉田卓彌, 上田泰孝(安藤ハザマ技研), 森則理, 俣野祐美(大建工業)

- ◆天井吸音を有する会議室の壁一面に乱反射性能が異なる3種の拡散体(Type-A, -B, -C)を設置した際の室内音響パラメータへの影響を時間領域FEMを用いて調査し、以下の知見を得た。
  - ・壁一面への拡散体の設置は、音場の高域における非拡散なふるまいを完全に排除できないが、室内の残響時間とEDTの低減、 $C_{50}$ の向上に寄与した。
  - ・残響時間は拡散体設置時も拡散音場理論値との乖離を示したが、EDTならびに $C_{50}$ については一面の拡散処理でも理論値への収束を示すケースがある。
  - ・各帯域で乱反射が最も低い拡散体が室内音響に与える影響が最も小さい。

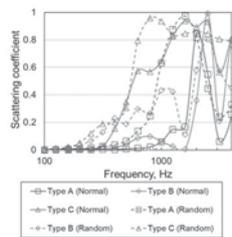


Fig.1 Normal- and random-incidence scattering coefficients of three 1D rib diffusers.

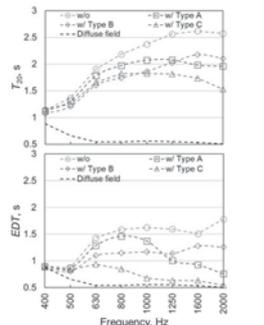


Fig.2 Acoustic parameters in the meeting room for the cases without diffuser and with diffusers:  $T_{20}$  (Upper) and EDT (Lower).

## 1-5-14

### 1-5-14 一般的なスピーカを用いた楽器の指向性模擬の試み —スピーカ配置の検討—

An attempt of simulating musical instrument directivities with a common loudspeaker: Loudspeaker arrangement

◎原量子, 岩見貴弘, 柿崎紘, 尾本章(九大・芸工)

- ◆楽器の指向性を考慮した室内インパルス応答を得ることを目的とし、一般的なスピーカを様々な方向に回転させ、球面調和領域で指向性を模擬する試みを行っている。
- ◆室内インパルス応答の収録における利便性を考慮したスピーカ配置を検討した。
- ◆次の3種類のスピーカ配置について指向性合成のシミュレーションを行い、各配置の模擬精度や実装に際しての利便性を比較した。
  - 等角度間隔の配置 (Step45)
  - 等密度性を考慮した配置 (Fibonacci)
  - スパースアルゴリズムを用いた配置 (OMP)



Fig. 1 : Loudspeaker arrangement of Step45

## 1-5-13

### 1-5-13 複合現実技術を用いた音響インテンシティ分布の時間変化の可視化 Mixed Reality Visualization for Sound Intensity with Time Variation

☆片岡優太, 及川靖広 (早大理工), 池田雄介 (東京電機大)

- ◆音の伝搬の可視化は、音響設計や音響教育に活用される。
  - 音響エネルギーが壁などの境界によって反射、回折といった影響を受ける様子を明らかにするため。
- ◆本稿では、複合現実技術を用いた時間変化を伴う音響インテンシティ分布の計測・可視化システムを提案した。
  - 測定は、音源の再生と録音を同期して、手持ちマイクロホンで音場内を走査することで容易に、かつ手短に行える。
  - 時間周波数領域の Pseudointensity vector (PIV) を可視化し、表示する分布図を時間フレームごとに変更することで、音のエネルギー伝搬を直感的に把握することが可能となる。
- ◆可視化実験を通して、スピーカから発する音がパーティションによって反射、回折する様子を確認できた (Fig.1)。

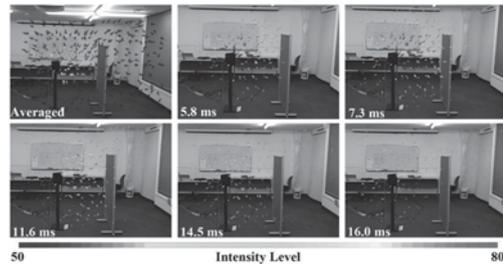


Fig.1: Distribution map of PIV. The top left figure shows the time averages, and the other figures show the intensity distribution for each frame.

## 1-5-15

### 1-5-15 FMBEMに基づく入射指向性解析 —その1 平面波展開に起因する誤差の検討—

Incidence directivity analysis based on FMBEM - part I: investigation on errors due to plane wave expansions

○安田洋介(神奈川大・工), 榎本貴之(サイバネット, 東大・新領域), 井上尚久(前橋工大), 佐久間哲哉(東大・工)

- ◆FMBEMに基づく入射指向性解析に関する基礎的検討として、3次元定常音場基本解の平面波展開を基に数値実験を行い、音場再現領域と入射指向性の関係や展開に関わる誤差要因について整理・考察した。
- ◆音場の平面波展開における展開係数は、展開点を含むある有限領域(音場再現領域)に対する入射指向性であると解釈できる。
- ◆展開の打ち切り次数を制御することで、所望の大きさの領域に対する入射指向性を算出できる可能性がある。

- ◆HF-FMBEMを用いて入射指向性を算出する場合、平面波展開の打ち切り次数  $N_p$  を HF-FMBEM の一般的な推奨値とした上で、低次の局所展開係数成分のみを用いて再構築すればよいものと考えられる。

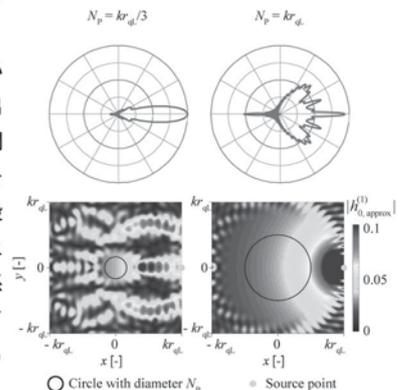


Fig.1: Incidence directivities and distributions of fundamental solution obtained using plane wave expansions.  $N_p$  denotes the truncation number.

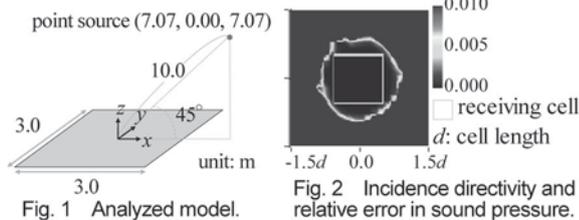
### 1-5-16

#### 1-5-16 FMBEMに基づく入射指向性解析 —その2 有限平板による散乱音場への適用—

Incidence directivity analysis based on FMBEM - part II: application to scattered sound field with a finite plate

○榎本貴之(サイバネット, 東大・新領域), 安田洋介(神奈川大・工), 井上尚久(前橋工大), 佐久間哲哉(東大・工)

- ◆実務的な問題を対象として入射指向性の計算方法をまとめ、球面波と有限平板による散乱音場における入射指向性を計算し、幾何学的な入射方向との対応や、対応する受信領域について確認した。
- ◆入射指向性の計算方法: FMBEM に基づく音場解析により、受信セルにおける内向係数を求める。この内向係数を、平面波展開の打ち切り次数に対して半分の次数までの局所展開係数から再構築する。
- ◆Fig. 1 に示す解析モデルで、音源の直接音が到来しない領域の受信セルにおいて、入射指向性と音圧の算出誤差(入射指向性から求めた音圧と、FMBEM に基づいて求めた音圧の相対誤差)を求めた(Fig. 2)。
- ◆受信セルから見た平板の両端部より、回折による音波が到来している。入射指向性に対応する受信領域は、受信セルの領域をカバーする。



### 1-5-17

#### 1-5-17 分散誤差を最適化した陽的時間領域 FEM による室内音響ソルバー

Room acoustic solve using a dispersion optimized explicit time-domain FEM

○吉田卓彌(安藤ハザマ技研, 神戸大院), 奥園健, 阪上公博(神戸大院)

- ◆時空間 4 次精度陽的時間領域有限 FEM (従来法) による室内音響ソルバーの有用性の更なる向上のため、分散誤差を最適化した手法を提案する。
- ◆分散誤差の最適化は次の手順で行われる。(1) 特定の音波伝搬方向における分散誤差を最小化するように要素行列の数値積分点を修正する。(2) 時間離散化の誤差が特定の周波数で最小となるように時間積分の重みを修正する。
- ◆提案法の広帯域の室内音響予測ならびに任意形状要素使用時の性能を剛境界をもつ室内音場解析を通して検証し、以下の結果を得た。(1) 提案法は、従来法に比べ、より高精度な広帯域の室内音場解析が可能である。(2) 提案法は直方体ならびに不整形要素を使用した場合も従来法より高精度に室内音場解析が可能である。

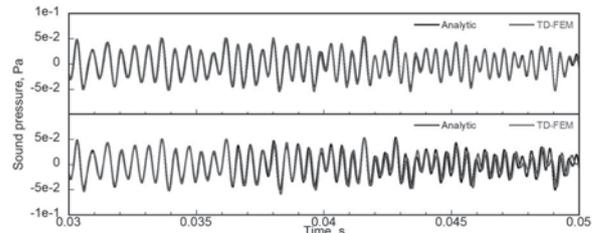


Fig.1 Comparisons of waveform between analytic solution and explicit TD-FEMs using rectangular elements: proposed method vs. analytic solution (Upper), conventional method vs. analytic solution (Lower).

### 1-5-18

#### 1-5-18 Partition of Unity FEM による室内音場解析 —従来 FEM との計算コストの比較—

Room acoustic modeling with Partition of Unity FEM -Comparison of computational cost with standard FEM-

○奥園 健, 阪上 公博(神戸大院・工)

- ◆本研究では、支配方程式の一般解を形状関数に導入することで場の近似を強化する Partition of Unity FEM による室内音場解析手法の確立を目指している。
- ◆本発表では、平面波エンリッチメントを用いて Helmholtz 方程式を解く Plane-wave-enriched FEM (PW-FEM) について、逐次計算ならびに OpenMP を用いた並列計算時の計算コストを従来 FEM と比較する。
- ◆数値実験を通して、以下を明らかにした。
- PW-FEM は逐次計算では従来 FEM に比べて、計算時間は長大となるが、必要メモリは大幅に少なく省メモリな手法である。
- 並列計算により大幅な高速化が可能であり、従来 FEM よりも計算時間が低減できる。
- 特に係数行列のスパース性を保持したまま自由度を著しく圧縮できるため、ロバストな直接解法をより高周波数まで利用できる点は、大きな長所である。

### 1-5-19

#### 1-5-19 Partition of Unity FEM による 3 次元室内音場解析に関する基礎的研究

Basic study on 3D room acoustics simulation using plane-wave-enriched FEM

☆田丸加奈子, 奥園健, 向江俊一, 阪上公博(神戸大院)

- ◆平面波エンリッチメントを用いた Partition of Unity FEM(PW-FEM)による 3 次元室内音場解析手法を提示する。
- ◆音響管内の平面波伝搬問題と立方体室の解析を通して、吸音境界を含む音場予測の精度検証を行う。
- ◆PW-FEM は波長より大きなサイズの要素による粗いメッシュを用いて、従来 FEM よりも少ない自由度で高精度に解析できることを明らかにした。

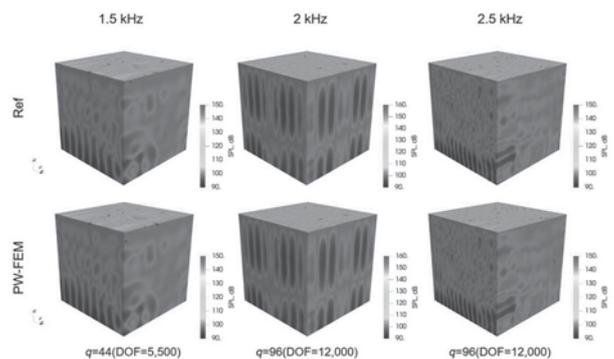


Fig.1: Comparison of SPL distributions between reference solution (Ref) and PW-FEM.

1-6-1

1-6-1 道路交通騒音予測において車両の音響パワーレベルの不確かさが予測値に及ぼす影響に関する検討

A Study on the Effects of Uncertainty of Vehicle Sound Power Level on the Road Traffic Noise Prediction

○井手上健斗, 山内勝也(九州大芸工)

- ◆道路交通騒音予測モデル "ASJ RTN-Model 2018" (ASJモデル) では、自動車走行騒音の音響パワーレベル  $L_{WA}$  をベースに予測を行うが、 $L_{WA}$  は車両や路面を要因としたばらつきを持つと考えられ、予測値に影響を及ぼすものと考えられる。
- ◆一方で、ASJモデルの音源式を定めているデータセットがどのような誤差性状を持っているのかは、明らかになっていない。
- ◆既報 [1] にて、計測における「不確かさ」の考え方を応用し、自動車走行騒音のパワーレベルに関する不確かさモデルを構築したうえで、密粒舗装の一般道を走行する小型車を対象とした分析結果を示した。
- ◆本発表では、排水性舗装および高機能舗装Ⅱ型の自動車専用道路を走行する小型車を対象とした分析の結果について報告する。

Table 1: Results of uncertainty analysis on vehicle sound power level (at expressway)

速度 $V$ [km/h]	$u_{total}$ [dB]		$u_{vehicle}$ [dB]		$\hat{u}_{road}$ [dB]	
	排水性	高Ⅱ型	排水性	高Ⅱ型	排水性	高Ⅱ型
$80 \leq V < 90$	2.46	1.92	1.54	1.66	1.92	0.95
$90 \leq V < 100$	2.53	1.75	1.56	1.53	1.99	0.86
$100 \leq V < 110$	2.52	1.90	1.48	1.78	2.04	0.66
$110 \leq V < 120$	2.59	1.61	1.41	1.51	2.17	0.55

[1] 井手上, 山内. 日本音響学会九州支部第14回学生のための研究発表会講演集. pp. 23-26.

1-6-3

1-6-3 道路交通ノイズマップ計算に関する基礎的検討

その3: 日本街区に適した受音点生成手法の提案

Fundamental Study on Calculation of Road Traffic Noise Map

Part III: A Proposal for three generation methods of receiving point for Japanese city blocks

◎原田和典, 平栗靖浩(近畿大), 大嶋佑也(新潟大),  
△齊藤由典, △跡部哲士(日本ミシュランタイヤ)

- ◆本報告では、日本街区における道路交通ノイズマップ計算に適した受音点生成手法を提案する。
- ◆これまでの検討から建物のファサードの中でもっとも騒音に曝露される箇所は建物角付近になることが多い。そこで建物角を確実に受音点とする受音点生成手法4種(Method0~3)を提案、CNOSSOS-EU手法2種、環境基準評価マニュアルをベースとした手法1種を含めた計7種の受音点生成手法について、より精緻に計算したものを真値とし、精度を検討した。
- ◆結果として建物角のみを受音点とする手法Method0が精度、計算コストともにCNOSSOS-EU手法を上回った。またそれ以外の3手法についても計算コストを許容できれば高い精度で計算可能であることがわかった。

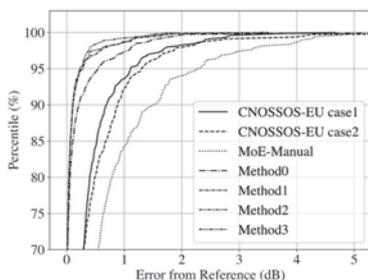


図 7 手法の誤差累積曲線

1-6-2

1-6-2 車両前後方向非対称の音源指向性を有する騒音予測モデルの特性に関する研究 -半無限長厚み障壁における ASJ Model との比較-

Study on characteristics of noise prediction model with asymmetric vehicle source directivity in front-rear direction: comparison with ASJ RTN-Model in semi-infinite length thick barrier

☆竹田郁哉(新潟大院・自然研), 大嶋佑也, 黒坂優美(新潟大・工)

- ◆前後非対称の指向性を有するHarmonoiseによる半無限長厚み障壁問題における等価騒音レベル分布を、ASJ Modelと比較した (Fig. 1)。
- ◆車両進行方向をx軸正の方向とした場合では障壁背後近く、進行方向をx軸負の方向とした場合では分布図左下で、両モデルの指向性補正時の等価騒音レベル差は最大となった。
- ◆両モデルの等価騒音レベル差は最大でも1.44 dBであったことから、中・大型車の交通量に占める割合が5%程度であれば、前後方向指向性の相違による等価騒音レベルへの影響は小さいと言える。

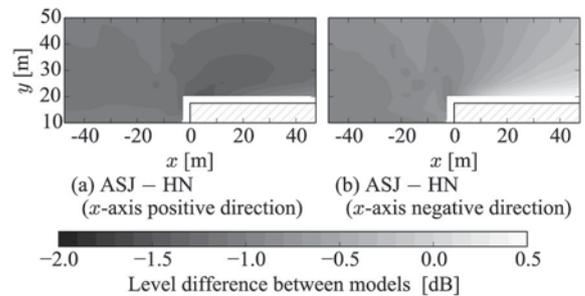


Fig.1: Distribution of the level difference between models with directivity (ASJ: ASJ Model, HN: Harmonoise).

1-6-4

航空写真を用いた道路交通騒音推定手法の実測による検証

Verification of road traffic noise estimation using aerial photographs by comparing with field measurement values

☆福田大輝(東大院), 米村美紀, 李孝珍, 坂本慎一(東大生研),  
平栗靖浩(近畿大)

- ◆広域の道路交通騒音推計のため、航空写真を用いて、断面交通量と走行速度を推定する手法を提案した。
- ◆推定したパラメータをもとに、道路交通騒音予測モデルの ASJ RTN-Model 2018 を用いて道路沿道の騒音レベルを推定し、同地点で行った10分間と30分間の短時間測定による実測値との比較を行った。
- ◆また、12時間の長時間測定による騒音レベルの実測値と、同測定で得た時間帯別の交通量と走行速度の計測値を用いた推定値、提案手法を用いた推定値との比較も同様に行った。
- ◆提案手法によるパラメータを用いた推定値は、国土交通省が調査・公表する道路交通センサスのパラメータを用いた推定値よりも、実測値との対応が高いことが認められた。

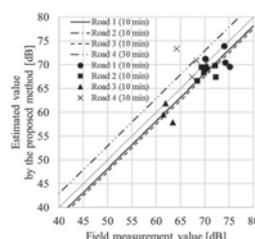


Fig.1 Comparison between the estimated noise levels by the proposed method and the field measurement results

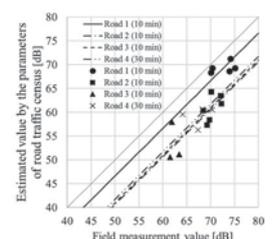


Fig.2 Comparison between the estimated noise levels based on the road traffic census and the field measurement results

### 1-6-5

#### 1-6-5 発進催促を目的とする警笛の使用に 影響を及ぼす要因の分析

Analysis of the factors of honking to urge the leading vehicle to start

☆三宅智也, 山内勝也, 稲田環, (九州大芸工),  
金基弘(駿河台大), 高田正幸(九州大芸工)

- ◆本研究は警笛使用に影響を与える要因の分析・検証を行い、法規定外の警笛使用を抑制するための提言につなげることを目的とする。本稿では、録画映像による警笛発生要因分析とドライビングシミュレータ(DS)を使用した実験を行った。
- ◆交差点での実測調査映像による分析から以下のことが示唆された。
  - 交差点での警笛は青信号点灯直後の発進催促が半数以上を占める。
  - 発進催促警笛は右折車線をはじめ現示変化に気づきにくい場合に多く発生する。
- ◆発進催促警笛が発生しやすい状況を再現しDS上で実験を行った。実験の結果、以下のことが示唆された。
  - 右折青信号のみ点灯する現示では運転者は信号の現示変化よりも対向車線の交通流に注目する。
  - 発進催促警笛の使用には「前方車両発進時間」、「進行方向」、「運転頻度」、「警笛に対する態度」、「発進催促警笛の使用に関する法律理解」が関連し、特に「発進催促警笛の使用に関する法律理解」が低い運転者ほど発進催促警笛を使用しやすい。
  - 周囲のネガティブな反応を引き起こしやすい「長めの警笛」は警笛を使用したいと思う場面が多い運転者ほど使用しやすい。
- ◆以上から、運転者の発進催促警笛に関する法律理解を高めていくことが警笛騒音緩和において重要かつ現実的方策の1つであるといえる。

### 1-6-7

#### 1-6-7 複数機が重畳した航空機騒音の ノイズネスに関する実験

An experiment on noisiness when multiple aircraft noises are overlapped

○森長誠(防衛基盤協会), 松井孝典, 桑野園子, 難波精一郎(大阪大学)

- ◆複数の航空機の運航において単発騒音が重畳した場合、「 $L_{A,Smx} - 10$  dB」の範囲では $L_{EA}$ の計算区間として不十分であることが報告されている。
- ◆このため「航空機騒音の測定・評価マニュアル」では、測定時に騒音の発生回数や機数の関係を記録することが望ましいとされているが、人間の聴覚情報だけで分離を行うことは容易ではないと思われる。
- ◆そこで、複数機が重畳する航空機騒音を刺激とした音響心理実験を行い、複数機の重畳を明確に判定するにはどの程度の時間差が必要であるか検討した。あわせて、重畳感の違いによるノイズネスの違いについても検討した。
- ◆Fig.1 に示す通り、聴覚情報だけで分離感が生じるにはかなりの時間差を必要としていることが示唆された。また、分離感の増加に伴いノイズネスが増加するという傾向は認められなかった。

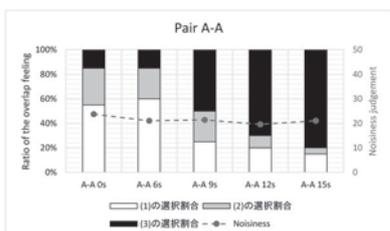


Fig.1: 2機の音源の時間差の違いによる分離感とノイズネス(結果の一例)  
横軸のキャプションは音源間の時間差を意味する。グラフの色が濃いほど、分離感が高いことを意味する。

### 1-6-6

#### 1-6-6 緊急事態宣言期間中の空港周辺における音環境の変化 —航空機騒音常時監視データを用いた予備的検討—

Preliminary study on change in environmental sound  
around an international airport during a state of emergency

○岡田恭明, 間瀬圭祐, 吉久光一(名城大・理工)

我が国では、空港周辺環境に及ぼす航空機騒音の影響は、時間帯補正等価騒音暴露レベル ( $L_{den}$ ) を用いて評価されている。また、各空港には複数の騒音監視測定局が常設され、長期にわたる  $L_{den}$  の状況が公開されている。今回、中部国際空港の2か所の常時監視局 (Fig.1, 空港までの距離: 約 11, 24 km) で観測された 10 分間隔の  $L_{Aeq,10min}$  と背景騒音より 10 dB 以上高い航空機騒音  $L_{A,Smx}$  (Fig.2) を用いて、就航便数が大幅に減少した緊急事態宣言期間中の音環境の変化について、過去 2 年間 (2018-19) の同時期のデータと比較しながら検討することとした。

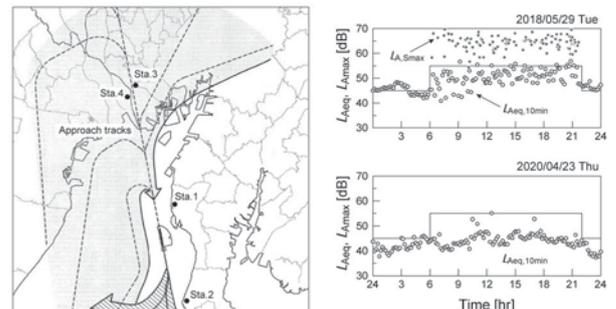


Fig.2 Examples of monitoring data at a station, Sta.4 in 2018 and 2020.

Fig.1 Typical flight tracks and monitoring stations around an intl. airport, NGO.

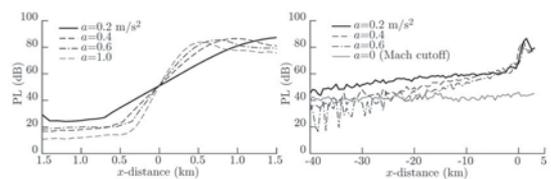
### 1-6-8

#### 1-6-8 大気擾乱を考慮したフォーカスブーム騒音の FDTD 解析

Finite difference-time domain analysis of focus boom noise  
with velocity disturbance in the atmosphere.

○土屋 隆生 (同志社大・理工/JAXA),  
金森正史 (JAXA・航空技術部門)

本稿では、次世代超音速旅客機 (SST) の水平加速に伴うフォーカスブームの解析を FDTD 法による 2 次元線形音波伝搬解析により試みている。ブームの波面位置について、音線追跡法の結果と FDTD 法の結果を比較したところ、良好な一致が見られた。また、大気中の速度擾乱の効果を考慮することで、Fig. のようにフォーカスブームの地表到達地点のはるか手前であってもブーム騒音の影響が発生することが示唆された。一方で、本手法は非線形効果を考慮していないことと数値分散誤差のために個々の波形を定量的に再現することは難しいが、全空間解析によりフォーカスブームのような複雑な音場の変化が容易に扱えるために、現象全体の構造を把握するには有用であると考えられる。



(a) PL distribution without velocity disturbance (b) PL distribution with velocity disturbance  
Fig.3: PL distribution on the ground surface at horizontal acceleration.

### 1-6-9

#### 1-6-9 音源移動に伴う音源群からの騒音伝搬性状の変化

Changes of the noise propagation characteristics from a sound source group due to sound source movement

☆牧野 裕介、高野 靖(京大院・工学研)

- ◆音源が移動すると放射音に周波数変調(Doppler効果)や振幅変調が起こる。直線軌道上を音源群が等速度運動した場合に、音源移動が騒音伝搬性状に及ぼす変化について検討した。
- ◆音源の移動によって振幅変調が起こると、受音点に接近する音源は振幅が増大し、受音点通過後の音源は振幅が減少する。
- ◆有限長線音源を点音源群に近似して騒音レベルを計算する(Fig.5)場合、近似する音源の個数の減少に伴って騒音レベルは増加する。この増加率は騒音の評価指標により異なる。(Fig.6)

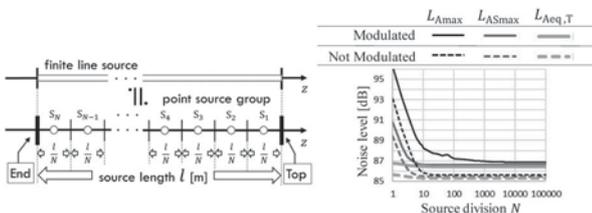


Fig.5: The model of finite line source and its calculation model

Fig.6: Noise metrics for different source division N: Calculated noise levels decrease as the source division increases.

### 1-6-11

#### 1-6-11 JIS規格に基づく聴覚保護具の遮音性能測定

The sound attenuation measurement of hearing protectors based on JIS standard

○横山 栄、小林知尋、土肥哲也(小林理研)、△傳田郁夫(聴力保護研)

作業環境における騒音暴露による聴力障害を防止するため、耳栓やイヤーマフ等の聴覚保護具の着用が不可欠である。2020年4月に改訂されたJIS T 8161-1:2020「聴覚保護具 第1部: 遮音値の主観的測定方法」に、聴覚閾値実験による遮音性能の測定方法が規定されているが、わが国では、旧JIS規格に準じた測定事例はいくつかあるが、現行JIS規格に準じた測定を実施している試験機関はまだ少ない。筆者らは、小林理学研究所の無響室内に現行JIS規格に準じた実験システムを構築し、イヤーマフを試験体として、遮音性能測定を試みた。旧JIS規格は製品規格で、耳栓2種、耳覆い1種の3種類に分類していたが、現行JIS規格によるSNR値、HML値等の指標を用いることによって、特定の騒音に対し、より適切な聴覚保護具の選択が可能となる。筆者らは、ANSI S3.19-1974に基づく遮音性能測定も実施しており、機会を改めて報告したい。

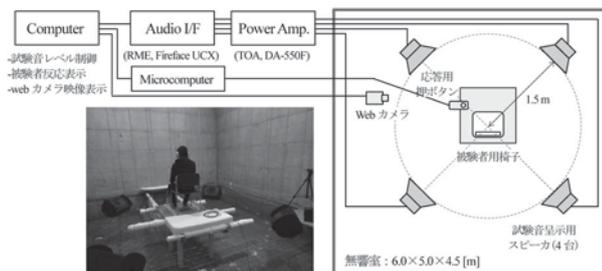


Fig.1: Experimental set up.

### 1-6-10

#### 1-6-10 聴覚保護具の遮音性能測定に関する規格の現状

Present standards for the sound attenuation measurement of hearing protectors

○横山 栄(小林理研)、山崎隆志(リオン/聴力保護研)

甚大な騒音環境における作業者の聴覚を保護するため、「騒音障害防止ガイドライン(平成4年)」では、適切な聴覚保護具の使用を求めている。イヤーマフや耳栓等の聴覚保護具の遮音性能測定法はJIS規格他に規定されている。本報では、主観的測定法に関するJISおよび各種国際規格の概要をまとめた。JIS改正の経緯をTable 1に示す。わが国では2020年、国際整合性を考慮して旧規格を廃止し、新たにISO 4869-1:2018を基としてJIS T 8161-1「聴覚保護具 第1部: 遮音値の主観的測定方法」が制定された。現状、各種国際規格でも、ランダム入射の1/3オクターブバンドノイズ(125~8k Hz)を用い、聴覚保護具装着時と非装着時の聴覚閾値の差を求め、これを遮音性能とする点は共通である。しかし、欧米における法規制の違いにより各国の適合規格は異なっており、ISO/TC43/SC1/WG17では次の改訂に向けた議論も始まっている。

Table 1 History of JIS standard for REAT

タイトル	規格番号	測定方法	評価方法
防音用耳せん	JIS B 9904:1955 (R 1958)	- pure tone (500-4k Hz)	耳栓: 1種, 2種
防音保護具 (耳栓, 耳覆い)	JIS T 8161:1974 (R 1978)	-1/3-OB noise (500-4k Hz) -frontal incidence -10 subjects, 3 trials	耳栓: EP-1, EP-2 耳覆い: EM ※製品規格
	JIS T 8161:1983	-1/3-OB noise (125-8k Hz) -random incidence	
聴覚保護具 (耳栓, イヤーマフ, 他)	JIS T 8161-1:2020 (対応国際規格 ISO 4869-1:2018)	- test site requirements - test subject qualifications -16 subjects, 1 trial - uncertainty of measurement	SNR値, HML値, オクターブバンド法 ※方法規格 JIS T 8161-2:2020

### 1-6-12

#### 1-6-12 ダミーヘッド(HATS)を用いたVHF領域の耳介周りの音響計測

Acoustic measurement around the pinna in the VHF region using a dummy head(HATS)

☆稲村 祐美、春澤 恒輝(神奈川工科大)、廣江 正明(小林理研)、長谷川 英之(富山大)、△神崎 晶(慶應大)、中村 健太郎(東工大)、上田 麻理(神奈川工科大)

- ◆本研究では、ダミーヘッドを使用し、頭の位置や向きが変化した際の耳介周りのVHF音を対象とした音響計測を行った。
- ◆1/4マイクロホンとMEMSマイクロホンを使用し、5つの周波数を呈示し、ダミーヘッドを5ピッチごとに回転させて収録した(Fig.1)。
- ◆FFT分析結果から呈示音の成分を抽出し、角度ごとにHATSなしを基準としたレベル差 $\Delta L$ を計算し、周波数ごとにレベルの平均と標準偏差を求めた。
- ◆分析結果から、周波数やマイクロホンの位置が変化すると $\Delta L$ の変化が生じることがわかった(Fig.2)。

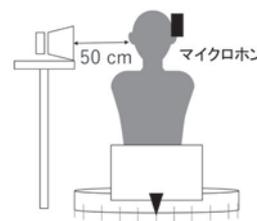


Fig.1 Position of microphone etc.

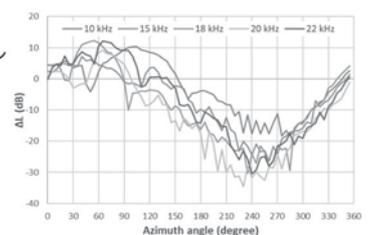


Fig.2 Relationships between azimuth angle and  $\Delta L$  measured by MEMS microphone.

### 1-6-13

1-6-13

#### 身の周りのVHF音その2

VHF sounds around us part-2

☆春澤恒輝、稲村祐美(神奈川工科大)、廣江正明(小林理研)、  
長谷川英之(富山大)、中村健太郎(東工大)、上田麻理(神奈川工科大)

- ◆ 先行研究において、家電から発生するVHF音の周波数成分を分析したため、そこでVHF音が発生していることを確認できた家電を中心に、音カメラを用いて音の発生箇所を特定し、様々な角度から収録して分析した。
- ◆ 結果として、LEDライトは電源部、サーキュレータは吸い込み口、吹き出し口から確認できた。周波数成分から見ると、LEDライトは2.5kHzピッチの高調波成分を有しているのに対し、サーキュレータは広帯域雑音を有しており、掃除機やドライヤーと特徴が似ていることから、両者のVHF音の違いはその発生要因と関係している可能性が高いということが判明した。



Fig.1: Measurement result of LED desk lighting (side)



Fig.2: measurement results of Circulator (side)

### 1-6-15

1-6-15

#### 畳下の床構造を対象とした 振動加速度推定フローの検討

Study on vibration acceleration estimation flow for floor structure under tatami

○富田隆太(日大・理工)、△後藤佑太(竹中工務店/元日大院・理工)、  
△青木怜依奈(日大院・理工)、足立大(リオン)

- ◆ 筆者らは、軟らかい防振ゴムを利用して、逆に低い周波数で防振ゴムと振動ピックアップによる設置共振を発生させ、その振動伝達率を利用することで、畳を剥がさずに、畳下部の床構造の振動を正確に測定しようとする方法について検討を行ってきた。
- ◆ 本報では、既報のカーペットに引き続き、加振器で得られた振動伝達レベルを利用して畳下の床構造を推定し、実測値との妥当性を検討した。その結果、建材量については、おおむね推定できることを示した。なお、本量については、今後の課題とする。また、畳下の床構造における振動加速度推定フロー案 (Fig. 1) を検討した。

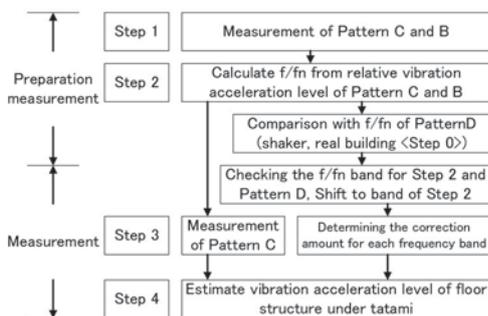


Fig. 1 Proposal of vibration acceleration estimation flow

### 1-6-14

1-6-14

#### 多孔質吸音材音響特性の確率的均質化解析

Probabilistic Homogenization Analysis of Acoustic Properties of Porous Sound Absorbing Materials

☆小松洋輔、山本崇史(工学院大)、

桂 大詞、遊川 秀幸、山川 啓介(マツダ)、立野 昌義(工学院大)

多孔質吸音材を対象として、均質化法へ確率論的手法を適用することにより微視構造の不確実性を考慮した吸音率の定量的解析手法を検討した。微視構造のばらつきを評価するため、ユニットセルの空孔率および相似変形を正規分布に従う確率変数とし、モーメント法にもとづく摂動法により吸音率の確率分布を予測した。求めた吸音率の確率分布をモンテカルロシミュレーションの結果と比較しその妥当性を確かめた。その結果、確率変数の変動が小さい場合には高い精度で吸音率の確率分布を予測できた。この時、摂動法の計算コストはモンテカルロシミュレーションと比較し大幅に削減された。

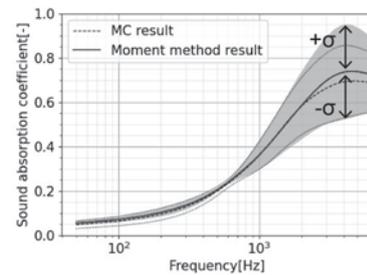


Fig. 1: Probability distribution of sound absorption coefficient for rectangular grid model. The porosity with average = 0.80[-],  $\sigma = 0.10$ , and the unitcell size with average = 100  $\mu\text{m}$ ,  $\sigma = 30$ .

### 1-6-16

#### 1-6-16 振動ふるい機の防音ハウス開口から放射される超低周波音に関する波動数値解析

—振動ふるい機のモデル化の影響と適切な位置の検討—

Wave-based numerical analysis of infrasound radiated from opening of soundproof house for vibrating screen machine: investigation on modeling and appropriate position of vibrating screen machine.

☆千田真人、関根秀久、安田洋介(神奈川大・工)、

川澄卓也、岩根康之、小林真人(飛鳥建設・技研)

- ◆ 振動ふるい機から発生する超低周波音のうち、防音ハウス開口から放射される成分に着目し、振動ふるい機の形状が音響放射パワーに及ぼす影響と、振動ふるい機の防音ハウス内での適切な位置について、周波数応答や音圧レベル分布を基に検討した。

- ◆ ふるい機の形状を変更しても放射パワーのピーク周波数の位置に大きな変化はなかった。防音ハウスの寸法は騒音の卓越周波数と上記ピーク周波数が重ならないようにすることで比較的容易に決定できる。

- ◆ 開口面に音圧の節が発生するようにふるい機の位置を調整することで、開口からの音響放射パワーを低減できる(本例では、位置を開口側に移動すると超低周波数域のディップ周波数が低域に移動した)。

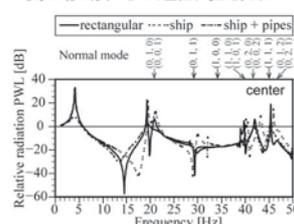


Fig. 1: Frequency responses of relative radiation PWL.

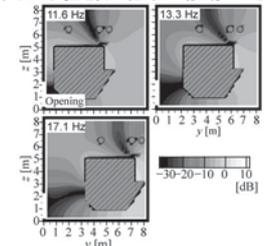


Fig. 2: Relative SPL distributions (ship + pipes).

1-6-17

1-6-17 居住床で発生する鉛直振動を対象とした振動感覚評価に振動知覚時間が及ぼす影響

Effect of vibration perception time on vibration sensation evaluation for vertical vibration generated in the residential floor

○松田貴(日大院・理工), 富田隆太(日大・理工)

- ◆「建築物の振動に関する居住性能評価規準・同解説」は2018年に改訂され、1/3oct.分析の加速度最大値を接線法のみで評価する方法から新たに継続時間の影響を加味して評価する方法に変更された。この継続時間を加味する方法として、時間補正係数を一定にしているが、時間補正係数は時間の経過とともに変化していくと考えられる。そこで、本報では、人の振動感覚に影響を及ぼす時間補正の更なる検討を行った結果を報告する。
- ◆検討の結果、衝撃振動、交通振動の両方において、振動知覚時間によって変化する  $kc_2' \times 10 \log_{10} Th$  が時間補正項として最も対応が良いことがわかった。

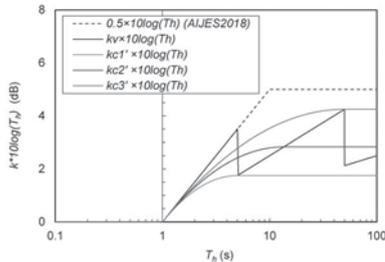


Fig. 1 Curve of time correction term  $k \times 10 \log_{10} Th$

1-6-19

1-6-19 ハイパースペクトル航空画像を用いた広域騒音伝搬予測のための地表面状態分類

Classification of ground states using airborne hyperspectral imagery for wide-area prediction of environmental noise propagation

○大嶋拓也(新潟大・工), 黒坂優美(新潟大院・自然研), 平栗靖浩(近畿大・建築)

- ◆ハイパースペクトル航空画像から、地表面の状態に応じた実効的流れ抵抗に対応する4クラスへの地表面分類を試みた。
- ◆トレーニングエリアとテストエリアの季節の組合せとして12パターンを検討した結果、いずれの季節の組合せでも、総合精度(overall accuracy: OA)の差は小さかった。
- ◆夏に撮影した航空画像による教師スペクトルを用いて、冬の撮影画像の推定エリアを対象として地表面分類を実施した(Fig. 1)。

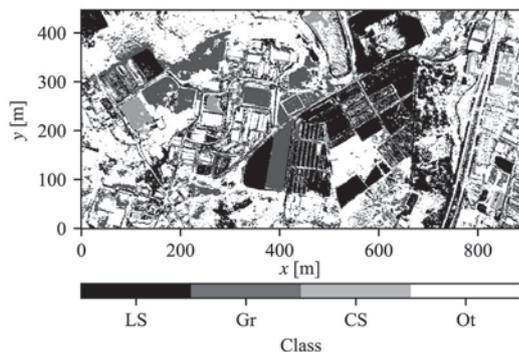


Fig. 1: Results of ground classification. LS: loose soil, Gr: grass, CS: compacted soil, Ot: others.

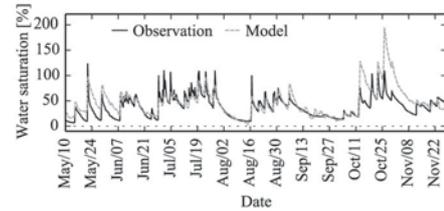
1-6-18

1-6-18 複数種地表面での長期測定に基づく地表面音響特性経時変化のモデル化に関する検討

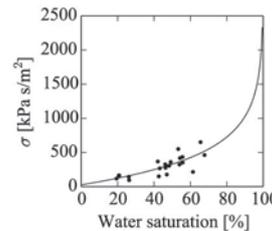
Study on modeling temporal variation of ground acoustic properties based on a long-term measurement campaign at various ground surface types

◎黒坂優美(新潟大院・自然研), 大嶋拓也(新潟大・工)

- ◆4種の地表面(畑, 草地, 田んぼ, 固い地面)において気象, 土壌水分量, 地表面音響特性の長期測定を行った。
- ◆気象要素から土壌水分量, 及び土壌水分量から地表面音響特性の経時変化のモデル化を, 既往研究 [Kurosaka, 2018] に倣い, 各地表面種につき試みた。
- ◆草地 (Figs. 1, 2), 畑, 田んぼでは, 経時変化モデルを概ね適用できた。



▲ Fig. 1: Observed and modeled hourly water saturations at grassland.



◀ Fig. 2: Scatter plots of the water saturation and the effective flow resistivity  $\sigma$ , and curve lines at grassland.

1-6-20

1-6-20 脳波による防災放送アナウンスの評価

Evaluation of the brain wave during the announcements for disaster prevention

○高野佐代子 土田義郎(金沢工大)

- ◆逃げる気になる防災放送の研究の一環として、脳波による反応による調査を行った。防災放送を模擬する弛緩刺激と緊張刺激を用い、得られた脳波から  $\alpha$  波と  $\beta$  波の各帯域の合計エネルギーから防災放送を聞いたときのリラククス度を比較することにより、防災放送における脳波を用いた研究の第一歩とする。
- ◆個人によってリラククス度の数値は違うが、全ての被験者において、緊張刺激の間はリラククス度が低かった。緊張刺激の防災放送は「脳波」という観点から見ても緊張していたと考えられる。

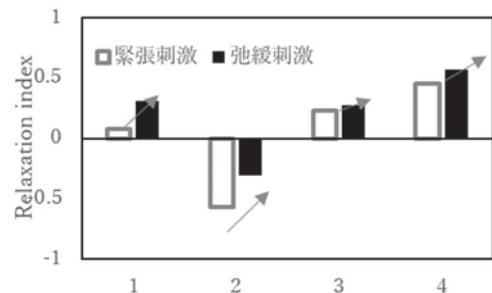


Figure 1. Relaxation index of each subjects for the tensed and relaxed stimuli

### 1-6-21

#### 1-6-21 地域連携による防災無線放送の音響的分析

Acoustic analysis of administrative radio system for disaster prevention with local cooperation

☆寺嶋将隆, △五十嵐康起(神奈川工科大), 三浦貴大(産総研),

中澤大洋(厚木市危機管理課), 小川喜道, 上田麻理(神奈川工科大)

◆昨今、多発している自然災害の発生時において、いち早く避難行動を行うために、迅速かつ正確な情報提供は必要不可欠であり、全国の自治体の多くが様々な防災システムの構築を喫緊の課題の1つとしており、本研究の最終目標としては、防災無線放送の情報が屋外であれば確実に伝わるようにすることである。

◆本研究では厚木市危機管理課全面協力の下、厚木市内16カ所の公民館職員を対象に防災無線放送の音声の聴こえの評価実験を実施した。また、並行して試験用防災無線放送(夕焼け小焼け)の長期観測を実施し、2020年12月26日現在まで100日分のデータを、スペクトログラムの画像データを基に分析を行った。

◆以上の結果を厚木市と神奈川工科大学地域災害ケア研究センター間で定期的に共有し、引き続き音声評価実験及び、長期観測を行っている。同時に、防災無線放送が聴こえにくいエリアのICTを用いた改善策についても模索している。

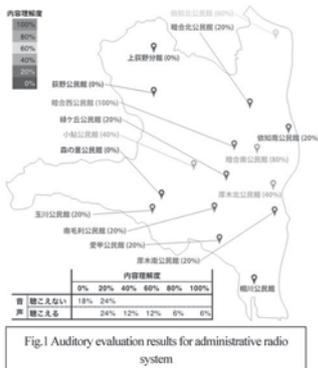


Fig.1 Auditory evaluation results for administrative radio system

### 1-6-22

#### 1-6-22 無形文化財保護のための騒音対策手法に関する研究 その2 -地域に根差した Noise Compatibility Planning の試み-

An investigation on the methods of noise countermeasures for safeguarding Intangible Cultural Property Part2 - Attempt of noise compatibility planning rooted in the local community - ○笠戸帆乃花(神奈川工科大), 廣江正明(小林理研),

△横岩康平(厚木市荻野地区自治会), △小川喜道, 田中哲雄, 上田麻理(神奈川工科大)

◆太鼓は古くから御祭りや儀式等で使用されてきた楽器であるが、近年、都市部でその演奏音は騒音と見なされることが多い。その為、地域文化活動やその継承活動の継続が危ぶまれる場合が少なくない。

◆本研究の最終目標は、こうした騒音問題の社会的対策による解決であるが、今回は、地域連携により騒音苦情の現状を把握し、その上で Noise Compatibility Planning の試みの一環として個々人の生活形態や状況に応じるべく太鼓の音に関する地域住民の声を収集するために、Web サイトを立ち上げ及び、運用システム構築を行った。

◆和太鼓の騒音から全ての人が逃れるわけではないが、小規模な地域ならではの個々人の特性等に配慮した騒音対策に重点を置いている。



Fig.1 Round meeting at KAIT



Fig.2 System configuration and How to use

### 1-7-1

#### 1-7-1 マイクロバブルを含む水流中における圧力波の非線形発展：非振動成分としての抗力と気泡並進運動が波に及ぼす影響の理論的発見

Nonlinear Evolution of Pressure Waves in Flowing Water Containing Bubbles: Theoretical Elucidation of an Effect of Translation of Bubbles and Drag Force Acting Bubbles as a Non-Oscillational Component on Waves

◎金川哲也(筑波大シス情), △谷田部貴大, △鮎貝崇広, △前田泰希(筑波大院シス情)

多数の球形気泡を一樣に含む水中において、各気泡と水が初期に等速で流れており、ここを伝わる平面進行圧力波の伝播を理論的に調べる(図1)。圧力波の無次元振幅は、有限ではあるが、十分小さいものとする(弱非線形音波)。気泡の生成・消滅・合体・分裂は考慮しない。

著者らのグループを含む、過去のあらゆる理論研究において、気泡に働く「抗力」と気泡の「並進」運動が無視されてきた。「抗力と並進」といった非振動成分は、圧力波という振動成分に影響を及ぼさないだろう」という合理的な先入観ゆえにではと推測する。本講演では、気泡の並進運動と、気泡に働く抗力に着目した理論解析を行う。その結果、低周波長波が従う KdV-Burgers 方程式および高周波短波が従う非線形 Schrödinger 方程式が導かれ、いずれの場合も、並進運動が波の非線形性を増加させること、抗力が波の散逸性を増加させることを示す。

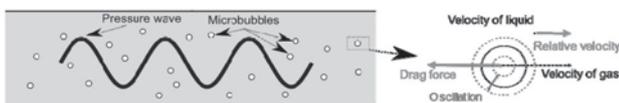


Fig.1: Pressure waves in flowing water containing many microbubbles (long wave case). Translation of bubble and drag force are incorporated.

### 1-7-2

#### 1-7-2 気泡含有水中の弱非線形音波は衝撃波とソリトンのどちらに発展するのか?

Whether Weakly Nonlinear Acoustic Waves in Bubbly Liquids Develop into Shock Waves or Solitary Waves?

◎金川哲也(筑波大シス情), △鮎貝崇広(筑波大院シス情)

多数の球形気泡を一樣に含む初期静止水中における平面圧力波の弱非線形伝播過程を記述する KdV-Burgers 方程式を、split-step Fourier 法を用いて数値的に解く。その結果、発展波形は、初期気泡半径が、0.1 mm 以下の場合には振動衝撃波に発展し、0.2 mm 以上の場合にはパルス波(音響ソリトン)に発展することがわかった(図1)。その一方で、初期気液体積分率は、発展波形の決定には概ね寄与せず、波の伝播速度に寄与することがわかった。

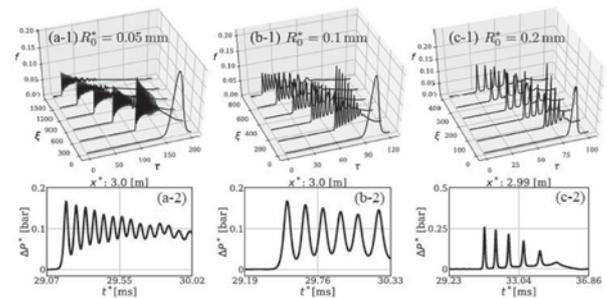


Fig.1: Dependency of spatial evolution of initially Gaussian waveform as 1D pressure waves in bubbly liquids on the initial bubble radius  $R_0$ . The initial volumetric fraction of gas phase is 0.01 [-], Initial pressure perturbation  $\Delta P^0$  is 0.48 [bar], time step is 1.0 [ms], and normal conditions of air-water system are used for computation.

### 1-7-3

#### 1-7-3 超音波照射によるマイクロバブルに吸着するDMPC分子の脱離量の定量的評価

Quantitative evaluation of the amount of DMPC molecules desorbed from a microbubble by ultrasonic irradiation

☆小林礼奈, 荻野瑞己, 田畑拓, 小山大介 (同志社大),  
Marie Pierre Krafft (CNRS)

- ◆超音波 DDS では、薬剤を付与したマイクロバブルを血流により運搬し、患部のみの特異的に吸着させ、超音波照射によりバブルを振動させ、薬物を患部のみを選択的に投与する。
- ◆超音波照射下によってバブルの周囲分子膜が受ける影響を検討するため、長距離顕微鏡と高速度カメラによる光学観測系によって、分子脱離量の定量的評価を行った。
- ◆ガラス基板上的バブル接触角と DMPC 溶液濃度の関係を測定した (Fig. 1(L))。本結果より、バブル接触角をバブル単位表面積当たりのモル濃度に換算することで、超音波照射による DMPC 分子脱離量を定量的に計測した (Fig. 1(R))。

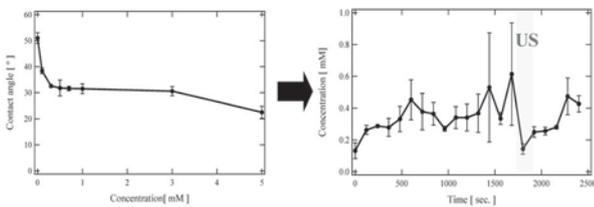


Fig.1(L) Relationship between the contact angle and the DMPC concentration and (R) change in the DMPC concentration on a bubble surface under pulsed ultrasound.

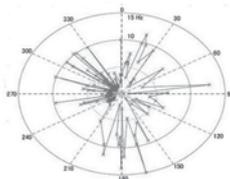
### 1-7-5

#### 1-7-5 活断層の放射特性における特異方位角

Internal structure and radiation waves of active fault model

○菊池年晃(防大), 水谷孝一(筑波大)

我々は、地震の活断層内の振動構造を音響学的知見から調べている。活断層の震動構造を解明するために、震源から放射された地震波を信号処理法を用いて解析している。複雑な地震波から基本振動を抽出するために特異値分解法を用いた。解析した特異値の大きさを調べた結果、初期の数項で特性を表すこと分かった。そこで級数展開した特異値の初項を震動の主成分と定義した。震源周囲全周にわたって主成分を求めた。今回は、2012年1月28日に富士山付近で発生したM5.4の地震に対して検証した。その結果を下図に示す。西側と東側で異なる性質の分布で構成されている。即ち、震動の放射波は方位的に特異双極性を示した。その成因を解明する為に、活断層をモデル化した幅の狭いダクトの放射伝搬のシミュレーションを行った。その結果、ダクトの前方と後方から放射されるパルスに大きな差異が生じた。即ちダクトの前方には衝撃性の急峻なパルスが放射された。一方、後方にはダクトの先端で反射した急峻なパルスがダクトを逆走する際に、ダクトによって濾過され緩やかな波に変換された。この波がダクトの後部から放射される。この現象を上図と対比すると、方位角 300° 方向には急峻な衝撃性パルスが放射される。一方、方位角 120° 方向には緩やかな形状のパルスが放射される。その結果、120° 方向を向いた等価音源が形成されたと考えられる。



### 1-7-4

#### 1-7-4 リピッドバブルを覆うリン脂質膜の粘弾的性質の実験的評価

Experimental evaluation of viscoelastic properties of phospholipid films coating lipid bubbles.

☆金兒千晶(千葉大・院融合), 章逸汀(千葉大・院医学薬学), 豊田太郎(東大・院総合文化), 林秀樹, 山口匡, 吉田憲司(千葉大・CFME)

- ◆異なる二種のリン脂質 (DSPC, POPC) を用いて造影剤を作成した。
- ◆DSPC および POPC はそれぞれ相転移温度が 58°Cと-3°Cであり、DSPC は室温で固体状、POPC は液体状の膜を持つと想定される。
- ◆バブル懸濁液の減衰特性の計測結果 (Fig. 1) から、DSPC バブルでは周波数が高くなるにつれて徐々に減衰係数が高値となり、POPC バブルでは低周波数側で高くなるという反対の傾向が確認された。
- ◆上記の違いは膜材質の状態の違いを反映していると考えられる。
- ◆減衰係数のピーク周波数を確認するために、周波数帯域を広げて検討する必要がある。

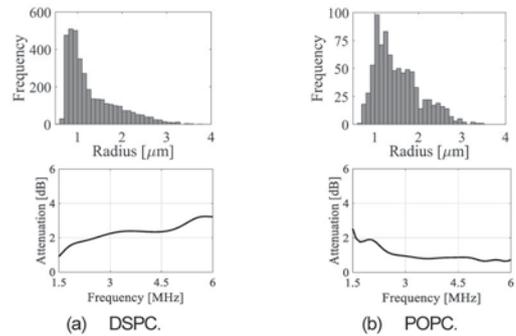


Fig.1: Radius distribution and attenuation of bubble suspension.

### 1-7-6

#### 1-7-6 等音速層の厚さに起因する音波の海底透過の変化

Variation in bottom transmission of sound wave caused by thickness of iso-velocity layer

○鶴ヶ谷芳昭 (山陽精工), 菊池年晃 (防衛大), 水谷孝一 (筑波大)

- ◆海底層が等音速の堆積物(ML)に覆われた場合、層の厚さによって水中の音場及び海底層への透過の状況は変化する。600 Hz の場合、MLの厚さが厚いと ML 上面からの反射が卓越する。また、1~0.5 m になると ML 上面からの反射より海底上面からの反射の影響が大となる。
- ◆ Fig. 1 に周波数 100 Hz、波長 15 m、音源深度 10 m、海底上の ML の厚さが 10 m の場合の音場(SF)を示す。a)は伝搬損失(TL)による音場を示し、b)は海底層のみによる音場を減算した減算音場を、c)は ML のみの海底とした音場を減算した音場を示している。b 図と c 図を比較すると ML 上面からの反射された音波の影響が卓越している。一方、海底下への透過は海底層を形成するシルト層と ML の物性値の違いによる角度の変化を生じる。また、ML の厚さが 6m 以下になると海底のシルト層の影響が増大する。さらに ML が薄くなるとシルト層の上面からの反射が増大する。海底への透過は透過パターン及び透過点は変化しないが、差音場のレベルは減少する。

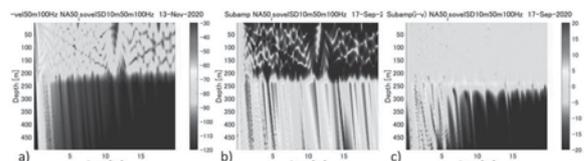


Fig.1: Comparison of sound field in thickness of ML; 3 m. freq.; 100Hz a) sound field (SF on TL) b) subtracted sound field [a] minus SF on silt layer bottom] c) subtracted sound field [a] minus SF on IVL bottom]

1-7-7

航行船舶の放射雑音測定のための海中設置ハイドロフォンに関する考察

Consideration of a hydrophone installed underwater to measure the noise of sailing vessels,

☆山田知佳, 清水悦郎(東京海洋大学院), 土屋利雄(JAMSTEC)

近年、海洋生物に対する船舶雑音の影響が懸念され、国際海事機関(IMO)は2014年に船舶による水中放射雑音低減のための非強制ガイドラインを採択し、今後の規制策的制定を目指している。我が国でも航行船舶の雑音計測が必須となると想定される。ISOによる雑音の標準計測法の簡易な方法として海底にハイドロフォンを設置して航行船舶の雑音計測を行う事が現実的である。2020年11月から数ヶ月間、大島沖にハイドロフォンを設置して船舶の雑音計測が実施されており、ここでは、実際の計測系を使った基礎的な音波伝搬シミュレーションを行い、Normal Mode法により伝搬距離と受波音圧レベルの関係を調べる。

◆ハイドロフォン設置海域とシミュレーション

船舶放射雑音計測は、ISO/DIS16554に準拠して行うのが一般的であり、計測船舶とハイドロフォンの水平距離を100m または船舶の長さ程度にする事が求められる。海中ハイドロフォンで計測する場合、AISの船舶位置情報から水平距離を求め、伝搬減衰を推定しなければならないため、シミュレーションから距離減衰特性を調査しておく必要がある。

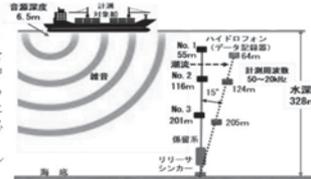


Fig.1 実海域における計測概念図

大島南方の海底に船舶放射雑音計測用のハイドロフォンをISO/DIS16554の計測手法に従いFig.1の様に設置し、計測を行う。潮流により係留系が15°傾いた場合も想定した。計測対象船舶は数千トンのディーゼル推進の貨物船を想定しプロペラ(深度6.5m)を雑音源とした。シミュレーションでは、音線法により伝搬経路を推定し、伝搬特性を求めるためにNormal Mode法によるシミュレーションを行った。Normal Mode法によるシミュレーションでは、対象周波数の50-200Hzの内の50, 100, 200Hzについて計算を行った。ここでは、扱う周波数が低いと高次のモードは生成されないと推定し、各周波数において生成される臨界角付近のモード数を求め50Hzでは5, 100Hzでは11, 200Hzでは14としている。

◆シミュレーション結果

Fig.2にシミュレーション結果の一例を示す。図の(1)の伝搬減衰パターンでは、赤い雑音レベルの減衰が小さく受波レベルが大きくなる。海面付近の船舶放射雑音は、600-700m程度まで概ね水平に伝播する。Fig.2(2)に各ハイドロフォン位置での伝搬減衰量を示す。この図から水平距離(R)100m付近までほぼ球面伝播特性(-20logR)を示し、それより遠方では円筒伝播(-10logR)よりも減衰が小さいことが見て取れる。実際の計測では、AISの船舶位置情報から水平距離を算出し、ハイドロフォンで取得された受波レベルとFig.2(2)の減衰特性から船舶放射雑音レベルの算定が可能と考えられる。しかし、実海域においてNormal Mode法を適用するために必要な精密な海底質の情報が得られないためモード数を変えて計算を行う必要があると考えられる。

◆今後の予定

大島沖での連続計測後、収録データの解析から求められた海上航行船舶の騒音特性は、船舶騒音の規制に関するIMOの国際会議等において発表される予定である。

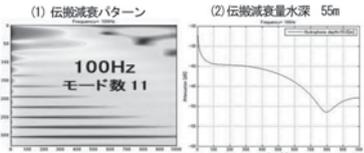


Fig.2 シミュレーション結果

1-7-8

1-7-8 周囲雑音イメージングに適用する凸型音響レンズの集束音場解析の予備的結果

Preliminary results analyzing sound field converged by a convex acoustic lens for ambient noise imaging

○森 和義, 田中 くるみ, 小笠原 英子(防衛大)

◆周囲雑音イメージングシステムをAUV艇首に搭載するための凸型レンズの集束音場を2次元FDTD法によって予備的に解析した。

◆解析領域には、海水、固体レンズ(シタクチックフォーム TG-28/4000)および内部液体(フロリナート FC-72)を配置し、線音源より500 kHzで5波のバースト波を放射した。各媒質の音速・密度・減衰定数は、水槽による実測値もしくは公称値を基にしている。ただし、固体レンズ内の減衰定数は実測によるばらつきを考慮し、低減衰で0.1 dB/λおよび高減衰で2.4 dB/λの2種類で設定した。(Fig.1)

◆入射角度0, 1, 2度におけるビームパターンは、入射角度ごとのメインローブが-3dBラインにおいて重複しておらず、本レンズの方位分解度は1度を実現することを示している。ただし、固体レンズの減衰定数の違いによって、集束音圧には約6 dBの差が生じている。(Fig.2)

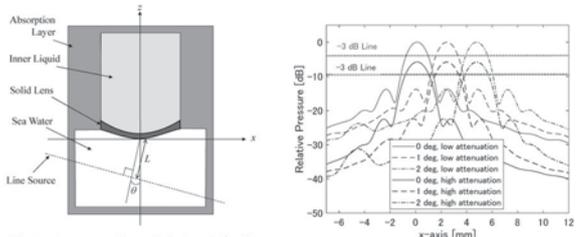


Fig.1 Arrangement in analysis domain for 2D FDTD method.

Fig.2 Comparison of beam patterns at 0, 1, and 2 deg

1-7-9

1-7-9 極浅海域における虚像音源を利用した時間反転波に関する実験

Experiment on time-reversal wave using a virtual image sound source in extremely shallow waters

☆篠田暲太, 太田和彦, 大川澄幸, 山本健介(金工大・工), 小川誠(NEC)

◆港湾内のような極浅海域ではマルチパス波が多く発生するため、一對の無指向性送波器においても、時間反転波を利用したパルスの収束効果が得られると考えられる。そのため、水深の浅い環境において、一對の無指向性送波器を使用して、パルスが収束するか水槽実験(Fig.1)を実施し、確認した。

◆結果として、一對の送波器において、パルスの収束(Fig.2)が確認された。また、受波器を上記の収束点から移動させると明確なパルス収束は確認されず(Fig.3)、元のパルスとの相関も弱くなった(Fig.4)。

◆また、パルスが収束する地点に水中物体が存在する場合(Fig.5)、元の1-パルス波と時間反転波をそれぞれ送波し受信エコーを比較すると、後者の方がエコーレベル0 dB(Fig.6 中一点鎖線)を上回る残響の頻度が少なかった。

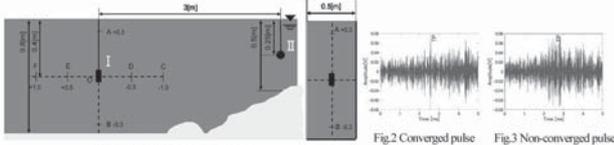


Fig.1 Arrangement of hydrophone

Fig.2 Converged pulse

Fig.3 Non-converged pulse

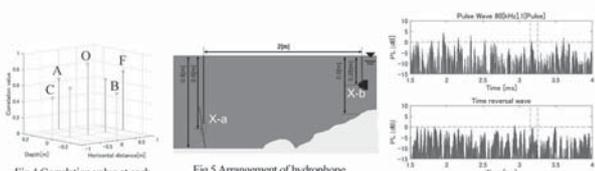


Fig.4 Correlation value at each receiving points (50 kHz)

Fig.5 Arrangement of hydrophone

Fig.6 Comparison of echo levels when 1-pulse wave (upper) and time-reversal wave (lower) were transmitted (0 dB: maximum of echo level)

1-7-10

1-7-10 ラット肝臓の音速評価における組織構造と試料作製法の影響の検討

Examination of effect of tissue structure and sample preparation conditions on speed of sound evaluation of rat livers.

☆瀬戸駿(千葉大・工), 野口和馬(千葉大・院融合工), 吉田憲司(千葉大学・CFME), 山口匡(千葉大学・CFME)

◆NASHモデルラット肝臓の凍結薄切試料を対象とし、中心周波数300 MHzの振動子を用いて音響特性解析を行った。

◆計測後の試料の音速像と病理像を比較し、凍結標本作製時の作業工程と組織構造が音速評価に与える影響を検討した。

◆スクロース濃度が高くなるにつれて、凍結時の氷晶現象による組織損傷が軽減されることが確認され、音速像には病理像に対応したテクスチャが示された。(Fig.1)

◆スクロース置換がない試料と比較して、スクロース濃度が高いほど音速が速くなることが確認されたが、その程度はパラフィン包埋試料における音速の上昇に比して小さかった。(Fig.2)

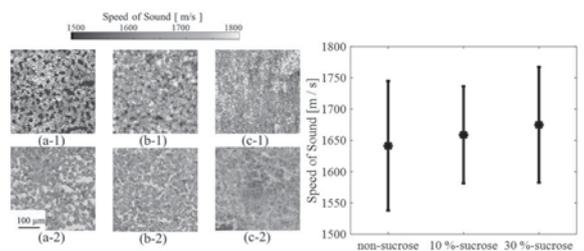


Fig.1: 2D speed of sound maps of sucrose-substituted NASH liver (a-1: non-sucrose, b-1: 10%, c-1: 30%), and histological images (a-2: non-sucrose, b-2: 10%, c-2: 30%).

Fig.2 Speed of sound of each sample with different sucrose concentration.

1-7-11

1-7-11 ラット肝臓音速のマイクロマクロサイズにおける周波数依存性の評価

Evaluation of frequency dependency of speed of sound of rat liver in from micro to macro size

☆伊能舞, 野口和馬(千葉大・院融合),

吉田憲司, 平田慎之介, 山口匡(千葉大・CFME)

- ◆臨床の周波数帯(1~15 MHz) および高周波帯(15~50 MHz)を用いた定量的超音波診断(QUS)に関する多くの研究が独立して報告されているが、固有音響特性と組織構造との対応付けがされていない。
- ◆生体組織を模倣したファントムやラット肝臓の音速を広い周波数帯域(1~250 MHz)で評価し、マクロサイズからマイクロサイズまでの組織構造と音速の関係性を検討した。
- ◆組織構造を有する肝臓において、15 MHz 以下では周波数依存性が認められないが、超高周波帯では強い依存性が確認された。
- ◆計測系の空間分解能が極めて高い場合には、細胞核や細胞質などの固有音速が評価可能であるが、臨床では高分解能といわれる15 MHz程度でも構造としての特性のみが評価可能であると思われる。

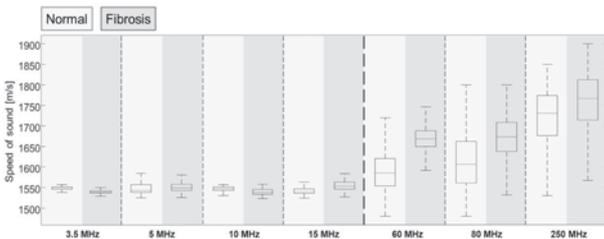


Fig. 1 Box-plots of speed of sound of normal and fibrosis rat livers.

1-7-13

1-7-13 NASH 肝臓の臨床エコーデータにおける振幅包絡評価モデルの適用性の基礎検討

Basic study on applicability of amplitude envelope statistical analysis model in clinical echo data of NASH liver

☆氏原裕貴(千葉大・工), 佐藤悠佑(千葉大・院融合),

田村和輝(浜松医科大), 森翔平(東北大院・工学研),

△Po-Hsiang Tsui(長庚大), 山口匡(千葉大・CFME)

- ◆脂肪量、線維化進行度が異なる症例を対象に、中心周波数 3.5 MHz の超音波診断装置で取得されたエコー信号について Multi-Rayleigh (MRA) モデルによる振幅包絡特性解析を行い、マイクロな組織構造と評価パラメータの関係性について検討した。
- ◆MRA モデルは媒質内に3種類の散乱体が分布し、振幅包絡分布に各散乱体からのエコー信号成分の足し合わせと仮定した。
- ◆線維化が進行した症例では評価パラメータ分布が動径方向に広がり、脂肪量が増加した症例では分布が中心にまとまるとともに偏角方向に広がる傾向が確認された。(Fig. 1(a), (b))
- ◆脂肪と線維が混在した症例では脂肪滴からのエコー信号が優位になり、脂肪肝と同様の傾向を示した。(Fig. 1(c))

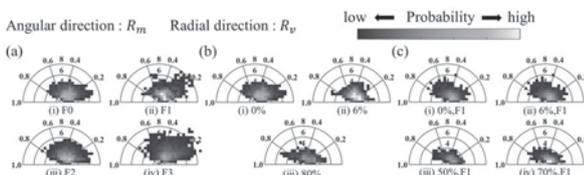


Fig. 1: MRA model parameters in polar coordinates for each clinical data: (a) liver fibrosis, (b) fatty liver and (c) NASH.

1-7-12

1-7-12 マウス担当リンパ節における散乱体構造と音速の関係性の基礎検討

Basic study of understanding of relationship between scatterer structure and speed of sound in mouse lymph node metastasis

☆野口和馬(千葉大・院融合), 大村眞明(富山大・工/千葉大・CFME)

小島拓也(東北大・工), 小玉哲也(東北大・院医工), 山口匡(千葉大・CFME)

- ◆マウスの腸骨下リンパ節に腫瘍細胞を播種後、経時的(0, 4, 6, 10 日)に *in vivo* でのリンパ節からの超音波エコー信号を取得し、振幅包絡分布を Nakagami モデルで評価することで散乱体分布を評価した。
- ◆摘出したリンパ節の薄切試料を対象に超音波顕微鏡計測および病理像化を行い、組織構造に対応した音速分布を評価した。
- ◆播種から日が進むにつれて Nakagami パラメータが低値を示す部位が多く確認され、腫瘍成長に伴って散乱体密度が低いまたは散乱体分布が不均一な部位が多く存在した。(Fig. 1)
- ◆組織構造の疎な箇所において音速が正常なリンパ節実質部に比して低値となることが確認され、腫瘍の成長に伴う組織構造の変化が超高周波超音波で評価される音響特性に影響を与えていることが示唆された。(Fig. 2)

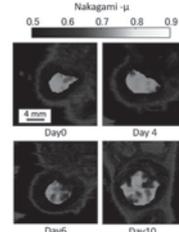


Fig. 1 Cross-sectional images of Nakagami- $\mu$  of metastatic lymph node.

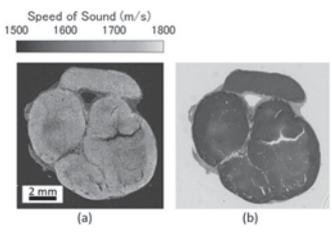


Fig. 2 Speed of sound map (a) and pathological image (b) of metastatic lymph node.

1-7-14

1-7-14 マウス筋芽細胞および HuH7 細胞の分化過程における音響インピーダンス変化

Evaluation of acoustic impedance in the differentiation process of mouse myoblasts and Huh-7 cells

☆大澤侑暉(千葉大・工), 橋本諒哉(千葉大・融合理工),

丸山紀史(順天堂大), 山口匡(千葉大・CFME)

- ◆マウス筋芽細胞と Huh-7 の分化過程による音響インピーダンス変化について検討した。
- ◆各細胞に対し、250 MHz の ZnO 振動子(方位分解能 7  $\mu$ m)を二次元走査しながら計測し、結果として得られた音響インピーダンスマップに適切な閾値を設定し、対象細胞の音響インピーダンスの平均値と標準偏差を評価した。経時変化を確認するため、この計測を5日間にわたって行った。
- ◆筋芽細胞の音響インピーダンスは細胞密度によっての変化はないが、いずれも経時的にわずかに低下する傾向が確認された。(Fig. 1(a))
- ◆Huh-7 においても細胞密度によっての音響インピーダンス変化は見られなかったが、経時変化によっての音響インピーダンス変化はほぼないことが確認された。(Fig. 1(b))

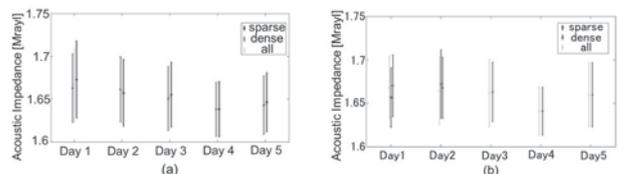


Fig. 1: Average of standard deviation of impedance of myoblasts (a) and average and standard deviation of Huh-7 cells (b)

## 1-7-15

1-7-15 頸動脈のビーム垂直方向における  
拍動に伴う変位と粥腫の変形の  
超音波弾性率計測への影響

Influence of displacement of carotid artery in perpendicular direction to ultrasonic beam and deformation of plaque on ultrasonic elasticity modulus measurement

☆土師佑太, 森 翔平, 荒川元孝(東北大),  
山岸俊夫(東北公済病院), 金井 浩(東北大)

- ◆本研究グループで提案されている血管壁の局所弾性率計測手法をアテローム性動脈硬化症において形成された粥腫の性状診断に用いることを目指しているが、弾性率の推定に用いる厚み変化の計測結果が、粥腫部位において、実際の頸動脈壁の動きに反する結果となってしまうという例が多くみられた。そこで、計測に影響すると考えられる、拍動に伴う超音波ビームに垂直な方向の血管の変位を計測し、その影響について考察した。
- ◆ブロックマッチング法を用いて健常な総頸動脈と頸動脈洞に形成された粥腫それぞれに対して長軸・短軸計測を行い、ラテラル方向変位を推定した。変位は長軸・短軸計測ともにラテラル方向の変位が健常部位と比較し粥腫部位で小さいという結果が得られた。
- ◆厚み変化計測の結果、粥腫の計測では、長軸・短軸計測ともに血管の拍動に伴う拡張により壁が薄くなるという実際の頸動脈の動きに反する結果が混在しており、健常部位の計測では実際の動きを反映した壁が薄くなる結果が全領域で得られた。
- ◆結果より、超音波ビームに垂直な方向の変位が大きいほど弾性率計測に対する影響が生じやすいわけではなく、粥腫が不均一性を有することが弾性率計測に大きな影響を及ぼしている可能性が考えられた。

## 1-7-17

1-7-17 単一の超音波プローブによる  
血流依存性血管拡張反応中の  
橈骨動脈の粘弾性特性変化の計測  
～パルス伝播時間法による血圧計測の導入～

Measurement of changes in viscoelasticity of radial artery during flow-mediated dilation using a single ultrasonic probe  
-Introduction of blood pressure correction based on pulse transit time method-

☆庄司悠人<sup>1</sup>, 森翔平<sup>2</sup>, 荒川元孝<sup>1,2</sup>, 大庭茂男<sup>2</sup>, 小林和人<sup>3</sup>, 金井浩<sup>2,1</sup>  
(<sup>1</sup> 東北大院・医工, <sup>2</sup> 東北大院・工, <sup>3</sup> 本多電子・研究開発部)

- ◆我々は動脈硬化症の早期診断を目指し、血流依存性血管拡張反応(FMD)中の血管の粘弾性計測を行ってきた。本報告ではパルス伝播時間法によって計測した血圧を用いて血管径の粘弾性を計測し、FMD中の粘弾性特性の変化を計測した。
- ◆20歳代前半の健常男性3名を対象に、橈骨動脈において単一の超音波プローブにより、FMD反応中の血管径の変化と血圧波形を同位置で同時に計測した。血圧波形はPTT法により推定した収縮期血圧と拡張期血圧によって1心拍ごとに校正した。
- ◆すべての被験者について駆血解放後に収縮期血圧と拡張期血圧がともに低下し、時間の経過に伴って安静時の値付近まで増加していく様子が確認できた。
- ◆弾性率は安静時においてばらつきはあるものの被験者間で同じオーダーで計測された。しかし、駆血解放後においては血圧の変化傾向は被験者間で同様であったにもかかわらず、弾性率の推定結果は被験者ごとに異なる傾向を示し、血管形状を考慮した粘弾性推定モデルの検討やFMDによる血管拡張を考慮した血圧推定が課題となった。

## 1-7-16

1-7-16 超音波後方散乱特性計測による  
赤血球集合体サイズ推定時の  
パワースペクトル算出条件の検討

Conditions for calculating the power spectrum for estimating size of red blood cell aggregates by analyzing ultrasonic backscattering characteristics

☆東山京平, 深瀬晶予, 森 翔平, 荒川元孝(東北大),  
八代 諭, 石垣 泰(岩手医大), 金井 浩(東北大)

- ◆非侵襲的かつ定量的な血液性状評価のため、超音波計測により求めた赤血球の散乱特性を単一球散乱体の理論特性と整合することで赤血球集合体サイズを推定する。本報告では、サイズ推定の際に用いる散乱パワースペクトルの算出条件について検討を行った。
- ◆単一球と仮定したワイヤ(径 10 μm)からの RF 信号  $x_0(n)$  とそれを  $\tau$  点遅らせた信号  $x_0(n+\tau)$  を足し合わせて干渉波形  $y(n, \tau)$  を作成し、各  $y(n, \tau)$  についてパワースペクトル  $P(f, \tau)$  を算出した。
- ◆各  $\tau$  について算出した  $P(f, \tau)$  を平均化することでサイズ推定に用いるパワースペクトル  $P_{ave}(f)$  を算出した。その際、干渉によりディップが発生するパワースペクトルの影響を低減するため、 $\tau = \tau_i$  における干渉波形  $y(n, \tau_i)$  の包絡線のピーク値  $p_{\tau_i}$  の二乗の値を重み関数として使用した。また、 $P_{ave}(f)$  を用いてサイズ推定を行った。
- ◆重み関数を使用することで平均化後のパワースペクトル  $P_{ave}(f)$  に対するディップの影響が低減され、サイズ推定をした際に重み関数を使用しない場合と比較して単一ワイヤの推定値(13 μm)に近い値に推定された。

## 1-7-18

1-7-18 超音波によって心臓収縮時の  
電氣的興奮が計測・可視化できる

○金井 浩, 荒川元孝, 森 翔平(東北大工)

- ◆大動脈弁の急激な閉鎖に伴って発生し心臓壁を伝播する機械的振動(横波)の速度は、心筋硬さ・壁厚に依存するため、伝播速度の超音波計測により心筋硬さ(横弾性率)が非侵襲推定できる。心II音付近で大動脈弁閉鎖に伴い生起する機械的振動伝播を用い、この方法適用の試みがある。
- ◆一方、心臓の収縮期には、僧帽弁閉鎖時に発生する心I音の機械的振動の伝播もあるが、前駆出期に電氣的興奮が房室結節からプルキンエ線維を伝導し、それに接する心室中等等の心筋が興奮・収縮を開始する。心筋虚血部では心筋収縮機能に異常が生じて伝播速度が低下することが知られ、前駆出期に心筋収縮の伝播を評価できれば心筋細胞の異常を早期検出できる可能性がある。
- ◆本報告では、超音波によって、心筋の興奮に伴う動きの速度成分を計測できることを示した上で、本グループが長年にわたり研究してきた位相差トラッキング法とパラレルビーム送受信により、心臓壁の貫壁方向と縦方向の2方向の速度を同時計測し、収縮応答の伝播速度が2方向で異なることが分かった。心臓縦方向と壁厚み方向で伝播速度が異なることは、中隔壁の螺旋運動を示す。このように収縮期では、機械的振動と収縮応答の成分が複雑に伝播していることが分かる。

### 1-7-19

#### 1-7-19 超音波バイオマーカーの確立に向けた 関連研究動向

Related research trends toward the establishment of ultrasonic biomarkers

○山口 匡(千葉大・CFME)

- ◆超音波診断は、消化器、循環器、産科領域を中心として広い臨床領域で実践されており、近年では定量診断法(QUS; quantitative ultrasound)も複数実用化されている。
- ◆一方で、信号や画像の特徴と生体組織の物理的性質および構造との対応が十分に取れていない。
- ◆日米の放射線および超音波医学系の学会が中心となり、各種の QUS パラメータの物理的意義を明らかにし、バイオマーカーを確立する検討が進められている。
- ◆関連して、QUS 法の新規開発や精度の向上、各種生体組織の音響特性の評価 (Fig. 1) なども行われており、工学基礎から臨床応用にわたって多様な研究サイトが連携した活動が活発な状況にある。

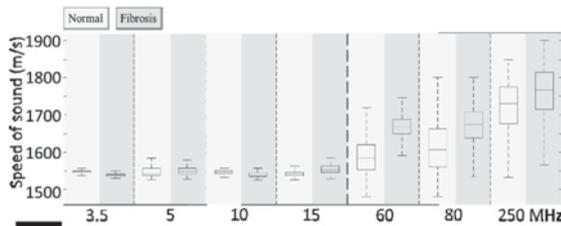


Fig.1: Speed of sound of rat livers evaluated from RF echoes observed with ultrasound of 3.5 - 250 MHz.

### 1-7-21

#### 1-7-21 超音波骨診断

-標準化と成長期の基準値設定を目指して-

Ultrasonic bone diagnosis -standardization and evaluation of bone growth-

○松川真美(同志社大), △尾崎悦子(京都府立医大)△真野 功(応用電機)

- ◆骨診断用定量的超音波法(QUS)の標準化を目指し、日本骨粗鬆症学会 QUS 標準化委員会では、初の試みとして踵骨用の国内 5 機種を対象に s-SOS (standardized Speed of Sound)と s-BUA (standardized Broadband Ultrasonic Attenuation)を提唱した。現在 2 機種がこの標準値を提示し、複数機関による横断的研究が実現可能となった。
- ◆近年の骨計測の流れは、皮質骨や海綿骨などの部位別診断にある。また原発性骨粗鬆症だけでなく、続発性骨粗鬆症の評価、子供の骨の成長の評価も視野に入ってきた。
- ◆超音波二波測定装置を用いて、十代の橈骨の皮質骨厚や海綿骨量の部位別大規模臨床計測を行った。皮質骨の成長が海綿骨より早いこと、また女子より男子の成長が早いことが確認された。しかし、十代後半でも若年成人の平均値に到達しない値もある。基準値を用いた骨の検診と、健やかな骨の成長のための健康指導が重要である。

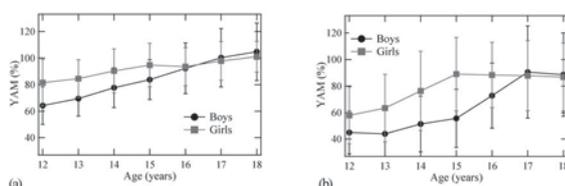


Fig.1: Changes of (a) cortical bone thickness and (b) cancellous bone density as a function of age. The data were normalized by the YAM (young adult mean) values.

### 1-7-20

#### 1-7-20 連続せん断波エラストグラフィによる 生体組織中を伝播するせん断波の可視化

○山越芳樹(群馬大)

生体軟部組織のせん断波伝播速度は癌等の疾病や運動器の弛緩緊張状態の違いで数倍以上も変わることがありこれがせん断波を医用診断に使う利点になっているが、伝播速度が大きく変わることは伝播速度の異なる境界で大きな屈折、回折、反射を生じ易く、このため、せん断波は生体組織内を複雑に伝播するようになる。

ここでは筆者らが考案した連続せん断波エラストグラフィ(C-SWE)によるせん断波の波面の可視化とその映像例を示す。これは体表に張り付けた小型のコイン型加振器でせん断波を励起し、組織内を伝播するせん断波の波面を汎用の超音波診断装置に付属するカラードプラ画像を用いて可視化する方法である。Fig.1 にこの方法で得た映像例を示す。筋組織、乳がん内のせん断波の伝播の様子が見て取れるが、せん断波伝播の可視化により伝播速度では得難い生体組織の弾性に関連した、例えば組織の癒着や映像面近傍の弾性構造の情報などが得られる可能性がある。

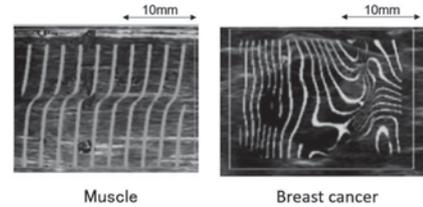


Fig.1 Shear wave propagated in tissues *in vivo*.

### 1-8-1

#### 1-8-1 英語発音評価における 文の発音難易度スコア抽出法の検討

Extraction of Sentence Pronunciation Difficulty on English Pronunciation Evaluation

☆内城 沙月, 伊藤 彰則, 能勢 隆(東北大)

- ◆本研究では、従来の発話単位の発音評価(発話スコア)から話者本来の発音能力のみ(話者スコア)を測るような評価システムを目指す。
- ◆話者の発音能力のみを測るために、直接話者の発音能力を推定するのではなく、まず発話している文の発音難易度(文スコア)を推定し最終的に話者の発音能力を測る方法が有効であると考えられる。
- ◆文の発音難易度推定を行うために、推定のための正解データと文の特徴量が必要であり、正解データは発話スコア分解モデルより抽出する。
- ◆本研究ではこれまでに行われていなかった発話スコア分解モデルについて検討を行った。
- ◆発話スコア分解モデルとして新しく Matrix-to-Vector Sum Decomposition (MVSD) を提案し、従来法である Non-negative Matrix Factorization (NMF) と比較を行った結果、MVSD の RMSE が低く、文の特徴量との相関も高いことから発話スコア分解モデルとしての MVSD の有効性が示された

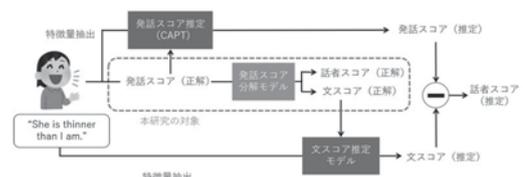


Fig.1: Block-diagram of a speaker's pronunciation ability evaluation

### 1-8-2

#### 1-8-2 IPAを介した音素-調音データ変換のための音素-IPA変換手法の検討

Phoneme-IPA conversion for generating articulatory movement sequences

☆飯山智晴, 大村英史, 桂田浩一(東京理科大学)

音声が生体内で共鳴する際には、舌や歯といった調音器官が動き、声道形状の調整が行われる。調音運動のデータ(調音データ)を用いることで、生成される母音の制御などが可能となる。先行研究では言語特徴量と呼ばれる特徴量を介して音素から調音データを生成していた。しかし言語特徴量を構成する調音情報は1音素からルールベースで生成されるため、正確とは限らない。また、言語特徴量の構成は研究者によって異なるため、システムの汎用性にも問題がある。そこで本研究では、言語特徴量の代替として国際音声記号(IPA)を用いて調音運動を生成することを目的とする。本稿ではそのための前処理として、音素列からIPA列を生成する手法を検討する。IPAは物理的な音を表すため、言語特徴量よりも正確に音を表現できる。また、国際音声学会が規格を定めているため、システムの汎用性の改善が期待できる。実験の結果、表1に示すようにニューラルネットワーク(NN)を用いた手法がモーラ変換ルールを用いた手法と比べて高精度なIPA列の生成を実現した。このことから、音素-IPA変換におけるNNの有効性を確認できた。

Table 1 Experimental results

	CER
提案手法(NN)	0.225
モーラ変換ルール	0.364

### 1-8-4

#### 1-8-4 段階的に制御された感情音声からの認知症者の感情認知に関する検討

A study of emotion recognition in gradually controlled emotional speech of people with dementia

◎齊藤千晶(認知症介護研), 中村篤(名市大院・理学)

- ◆我が国の認知症高齢者数は、高齢化率の上昇とともに増加すると予測される。
- ◆認知症症状の進行から、他者とのコミュニケーションや周囲の状況把握に困難が生じさせ、社会生活を送る上で支障を与えるため、認知症者との円滑なコミュニケーションについて考えることは重要であると考えられる。
- ◆今回、感情表現の強さに着目し、健康高齢者と認知症者を対象として、音声による感情表現に対する認知症者の理解の特徴について客観的知見を得ることを目指す。
- ◆認知症者は健康高齢者よりも有意に感情理解の正答率が低下することが示された。これは、感情強度が原音声と同等であるレベルから生じていた。さらに、感情表現の違いからも正答率に有意な低下が認められた。
- ◆認知症者とのコミュニケーションにおいて、感情を伝える際、留意する必要があると考えられる。

### 1-8-3

#### 音韻象徴性における調音ダイナミズムの研究 -空間認知と調音運動軌跡の関連性について-

The Articulatory Dynamism of Sound Symbolism: An Examination of the Relationship between Spatial Cognition and Articulatory Movement

○横山安紀子(日本大学)

本発表では、音韻象徴性に関する無意味語実験から発表者が考案した「調音ダイナミズム」モデルの検証を行った結果を報告する。ここでは被験者を日本語母語話者として実験を行った。これは英語母語話者に行った実験の追試である。実験は音から音への調音点移動において、異なる距離を表すCVCからなる無意味語を聴覚刺激として、また異なる空間距離を表す図を視覚刺激として組合せを作り、マッチングの検証を行った。全体の結果として否定的回答が肯定的回答を上回り、多数見られたが、中でも被験者が調音点の移動軌跡と空間距離の「大-小」が相反する組合せについて否定的回答が多く見られた。

英語母語話者による前実験結果においても、相反する組合せに対しては否定的判断をする傾向が強く見られたことから、ここでの結果が前実験の結果をある程度まで裏付けるものとなった。特に、図1に示すように、調音点移動距離と空間距離の相反する組合せを否定する回答が3次元の空間表現(thin, thick)に多く見られた。

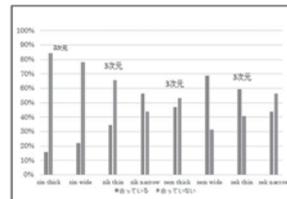


図1. 空間距離特徴を異にする無意味語と意味概念の組合せにおける肯定的および否定的回答率

### 1-8-5

#### 1-8-5 バブルノイズ環境下におけるポップアウトボイスの評価実験

Evaluation experiment of pop-out voices in babble noise

○天野成昭(愛知淑徳大), 河原英紀(和歌山大), 坂野秀樹(名城大), 牧勝弘(愛知淑徳大), 山川仁子(尚綱大)

背景雑音から際立って目立つ音声をポップアウトボイスと呼ぶ。この音声の性質を解明するための第一歩として、男女各5名の話者の音声を重畳して得たバブルノイズの背景雑音下において780話者(男性388名, 女性392名)の音声を12名の実験参加者にヘッドホンで呈示し、各話者の音声背景雑音からポップアウトする程度を5段階で評価させた。その結果をFig. 1に示す。Fig. 1から様々なポップアウトの度合いの音声幅広く存在することが分かる。また、少数ではあるがポップアウトの度合いが極めて高い音声が存在することが見て取れる。今後、各音声の基本周波数、スペクトル形状、振幅包絡の変動パターン等の音響特徴と、この評価データとの関連を解析し、ポップアウトボイスの性質を解明する予定である。

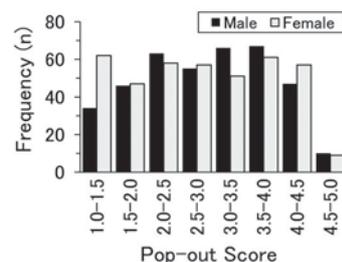


Fig. 1. Distribution of pop-out score of speech items in babble noise.

### 1-8-6

1-8-6

#### ポップアウトボイスの音響解析

Acoustic analysis of pop-out voice

○坂野秀樹(名城大), 牧勝弘(愛知淑徳大), 河原英紀(和歌山大),  
山川仁子(尚綱大), 天野成昭(愛知淑徳大学)

- ◆ポップアウトボイス(背景雑音から際立って目立つ音声)の音響解析を行った。
- ◆幅広いデータベースから選んだ790刺激を分析
  - 各刺激音にポップアウトの程度を表す主観評価値を付与済み
- ◆下記の音響特徴と主観評価値との間に一定の相関が認められた
  - 静的な情報を表す特徴量: 1kHz以上の帯域のパワー(Fig.1), 500Hz以上のスペクトル形状, 基本周波数の平均値, A特性によるRMS値
  - 動的な情報を表す特徴量: デルタケプストラムのノルム(Fig.2)
- ◆この解析結果から, ポップアウトボイスの特徴には, 静的な音響特徴と動的な音響特徴の両者が複雑に絡み合っていると考えられる。

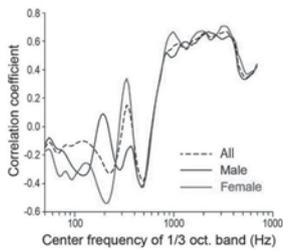


Fig.1: Correlation coefficient between the subjective score and the RMS value of each frequency band.

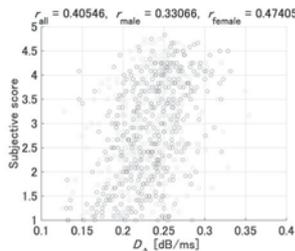


Fig.2: Scatter plot of the subjective score versus the norm of delta-cepstrum.

### 1-8-8

#### 1-8-8 rtMRI 動画を用いたオペラ歌唱における呼気の制御に関する研究

Research on the control of exhalation in opera singing using rtMRI

☆戸田菜月, △工藤理佐子, 竹本浩典(千葉工大), 高橋純(大阪芸大)

- ◆オペラ歌唱に特有の安定した豊かな声量は, 呼吸に関する部位の制御を高度化することにより獲得されると考えられており, 前報では横隔膜の上下方向の変位に基づいて, 歌唱中の呼気の制御を検討した。
- ◆本研究ではより高い精度で呼気を検討するため, 3人のオペラ歌手を被験者とし, 肺の形状の時間変化を分析した。また, 肺の形状変化を主成分分析し, 個人差や歌唱技術との関係を検討した。
- ◆肺の面積の変化と横隔膜の上下方向の変位を比較検討したところ, 肺の面積の変化は横隔膜の上下動から推定できることを示していた。また, 肺の面積の変化から呼気流量の推定することは困難であり, 冠状断面での面積変化を検討する必要があることも明らかとなった。
- ◆主成分分析を行った結果, 肺の形状変化のPC1, PC2を構成する運動要素と寄与率は個人差が大きかった(Fig.1)。また, 高音から低音の移行時に用いられる「支え直す」歌唱技術は, 胸部をほとんど動かさず, 横隔膜を一時的に下降させる運動であると考えられた。

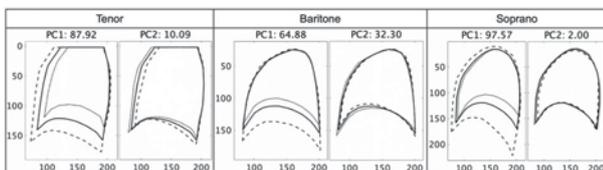


Fig.1: Results of principal component analysis of lung shape during singing (bold: mean, dashed: maximum score, dotted: minimum score)

### 1-8-7

#### 1-8-7 残響環境では無声破裂音がより有声音に知覚される現象について

Voiceless plosives tend to be perceived as voiced in reverberant environment

○荒井隆行, 岩上純也, 大澤恵里(上智大・理工)

- ◆長い残響環境において音声聴取が難しくなるが, その原因として, self-masking や overlap-masking が論じられてきている。
- ◆子音でも, 時間と共に音響的特徴が変化する破裂音では特に self-masking が生じやすく, 残響の影響を受けやすいと考えられる。
- ◆日本語破裂音の有声性に対する音響関連量としてのVOT(voice onset time)も, 残響がかかると推定が難しくなる。
- ◆そこで, VOTを変化させた日本語合成音声「タ」・「ダ」の連続体に対して模擬残響環境下における識別実験を実施した。



- ◆その結果, 「ダ」回答率が50%になる有声・無声の境界が, 残響がない条件でVOT = 10ms前後であったのに対し, 残響条件ではVOT = 20-25msへとシフトした。
- ◆聴取者にとって, 残響がかかった場合に破裂音のburstは比較的容易にその位置を把握できる一方, 有声開始の位置を判断することは難しかったものと考えられる。
- ◆判断が難しい場合に予測が行われ, その際に hyper-correction が働いた可能性が示唆された。

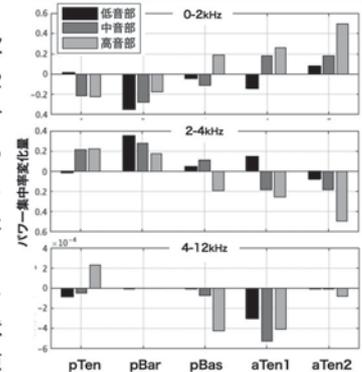
### 1-8-9

#### 1-8-9 オペラ歌手が非声楽的な発声を模擬した際の歌声の音響的特徴

Acoustic Analysis of singing voice when opera singers imitate non-operatic singing style.

◎高橋純(大阪芸大), 戸田菜月, 工藤理佐子, 竹本浩典(千葉工大)

- ◆優れた歌手の歌声には, 歌い手のフォルマント(Singer's formant(以下SF))と呼ばれる音響的特徴が存在する。
- ◆本研究では, オペラ歌手が熟達した歌唱法を用いて発声(声楽的発声)した際と, その歌唱法を用いずに発声(非声楽的発声)した際の声道形状, 歌声の音響的特徴, 歌声の評価結果を比較検討した。
- ◆0.2kHz, 2.4kHz, 4-12kHzの各帯域で, 声楽的・非声楽的発声のパワー集中度の差を検討した結果, プロと学生でスペクトル形状の変化のパターンが異なり, プロのパワー集中度は, 低次フォルマントを含む0.2kHzで減少し, SFを含む2.4kHzでは増加するが, 学生では逆になることが明らかとなった(Fig.1)。



- ◆また, rtMRI動画によって歌唱時の声道の形状を観察した結果, その原因が喉頭の位置や咽頭

Fig.1: 各周波数帯域のパワー集中度の変化の容積の変化であることが示唆された。

### 1-8-10

#### 1-8-10 ドイツ歌曲における声楽経験とドイツ語習熟度にかかわる音響特徴量の分析

An analysis of acoustic features related to singing and German-language proficiency for German songs

◎粕谷麻里乃(東邦音大・音), 荒井隆行(上智大学)

◆目的

本研究は、歌唱において美しい響きを担う母音を対象に、歌唱上の「発声」と「発音」の実態を明らかにする。

◆手法

声楽経験やドイツ語習熟度を考慮して集められた被験者を対象とする。先行研究において提案された歌声の習熟度を評価する SFR や 2 凸を指標として母音部分の音響特徴量を定量化し、「発声」を評価する。フォルマント周波数 (F1/F2) の分布により、「発音」を評価する。

◆結果

Fig.1のSFR/2凸, Fig.2のF1/F2において、対象母音(/i, /e/, /a/, /o/, /u/)に習熟度の影響を確認した。弱音節箇所 Singer's formant 周辺の周波数が顕著であることを確認した。各母音のF1/F2の統計検定により、F1/F2が声楽経験かドイツ語経験のいずれかの影響を受けやすいことを確認した。

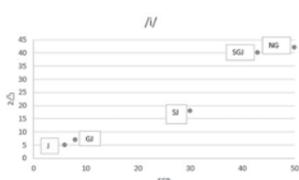


Fig.1: SFR/2凸 of /i/. (5 Participants NG, SGJ, SJ, GJ, J)

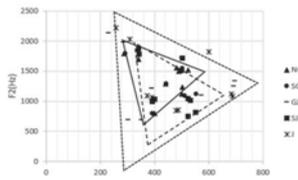


Fig.2: F1/F2 of each vowels /i, /e/, /a/, /o/, /u/. (5 Participants NG, SGJ, SJ, GJ, J)

### 1-8-12

#### 1-8-12 日本語話者 20 名の rtMRI 動画における発話器官の輪郭抽出

Speech organ contour extraction from rtMRI videos for 20 Japanese subjects

☆天野沢海, 並木崇宏, 宮川翔多, 後藤翼, 竹本浩典(千葉工大), 北村達也(甲南大), 能田由紀子, 前川喜久雄(国語研)

◆日本語の調音音声学の精緻化のために、発話運動のリアルタイム MRI 動画データベースを構築し、発話器官の輪郭を機械学習により点群として抽出する研究を進めてきた。

◆本研究では、以下の手順で 20 名の動画から発話器官の 5 つの部位ごとに輪郭を抽出した。

- 機械学習で輪郭抽出を行う矩形領域のサイズと輪郭線の終点・始点の座標をクラスタ分析し、特徴的な 8 名を選出した。
- 8 名それぞれに対して、各部位の形状のパリエーションを網羅する動画フレームから輪郭をトレースし、学習器を生成した。
- 生成した学習器を用いて、全被験者の動画の全フレームから輪郭を自動抽出した。

◆学習に用いていない 12 名の動画に対して輪郭抽出の精度を定量的・定性的に評価した結果、喉頭蓋・声帯を除いて良好であった (Table 1)。

Table 1 輪郭抽出精度の5段階評価 (5: 最高, 1:最低)

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	F1	F2	F3	F4	F5	F6	F7
口唇・下顎	4	4	4	4	3	-	3	-	5	3	-	4	-	-	4	4	-	-	-	4
舌	4	4	-	-	4	5	4	3	4	-	-	4	-	4	3	-	-	-	-	4
軟・硬口蓋	3	2	-	-	5	-	5	3	5	4	-	3	-	3	4	5	-	-	-	2
咽頭後壁	-	4	-	-	4	5	4	4	-	4	-	5	2	-	5	-	-	-	4	5
喉頭蓋・声帯	5	-	-	5	2	4	5	2	5	5	-	-	-	4	1	-	-	-	1	3

### 1-8-11

#### 1-8-11 内視鏡下鼻副鼻腔手術による術前・術後の形状と音響特性の変化の検討

Examination of changes in shape and acoustic characteristics of before and after operation by Endoscopic Sinus Surgery

☆杉浦唯, 竹本浩典(千葉工大), 北村達也(甲南大), 鴻信義(慈恵医大)

◆鼻腔・副鼻腔は手術によって形状が変化し、音声の個人性にも変化が生じる。

◆鼻音生成時の術前・術後の鼻腔・副鼻腔の形状をCTで計測し、その形状と音響特性の変化をシミュレーションで検討することによって、手術による音声の変化を予測しようとしている。

◆本研究では、内視鏡下鼻副鼻腔手術を受けた被験者 1 名の術前・術後の形状変化と音響特性の変化を検討した。

◆ その結果、全副鼻腔を開放したことにより、ディップ周波数とディップ生成に関与する分岐管の組み合わせが変化し、特に上顎洞と前頭洞が音響的に一体化する傾向が強くなった。

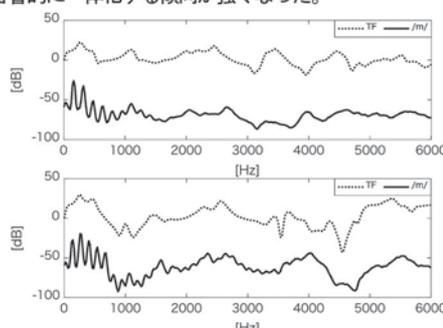


Fig.1: Spectrum (solid line) of /m/ and transfer function (dotted line). (Upper : preoperation, Lower : postoperation)

### 1-8-13

#### 1-8-13 EMA による調音運動観測のための咬合面と口蓋の計測法

Measurement of occlusal plane and palate shape for observation of articulatory movements using EMA

○能田由紀子(国語研), 北村 達也(甲南大), 浅井 拓也(早稲田大), 竹本 浩典(千葉工大), 前川 喜久雄(国語研)

◆EMAによる調音運動の観測では調音器官上に設置したセンサ位置だけでなく、咬合面や口蓋の正中矢状断面の形状が必要である。

◆歯科用印象トレイに 4 つのセンサを付加して実験参加者の上顎の歯列と口蓋の印象を採取し、同時に EMA によるセンサ位置を取得した。

◆取得した印象を 3D スキャナで計測し、4 つのセンサの座標をスキャンデータから取得することにより、スキャナで計測した口蓋形状を EMA の座標系へ写像する変換行列を求めた。

◆印象の左右の切歯および大臼歯上の 4 つの計測点から咬合面と正中矢状面を決定し、4 回の計測による口蓋上の正中矢状断面の位置のばらつきは、側方向に最大で 1.1 mm であった (Fig. 1)。

◆口蓋の正中矢状断面の形状を EMA の座標系へ移動し、安静時における調音器官上のセンサとの位置を検討した (Fig. 2)。

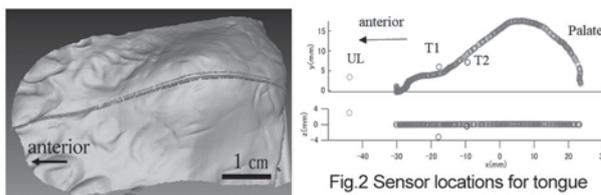


Fig.1 3D scanned palate shape and 4 midsagittal lines calculated from 4 data sets of teeth position

Fig.2 Sensor locations for tongue (T1, T2), and upper lip (UL) relative to the midsagittal palate shape during rest position. Upper: lateral view; lower: top view.

### 1-8-14

#### 1-8-14 real-time MRI 動画をを用いた音声合成システムの作成

Speech Synthesis from real-time MRI videos

☆丹治涼, 大村英史, 桂田浩一(東京理科大)

発話中の舌の動き等, 音声生成に関わる調音器官の動作を調音運動という。調音運動から音声を直接出力するシステムは言語情報を用いずに音響パラメータを推定できるため、後天的な音声障害者のための自然な代用音声への応用が期待できる。調音運動の観測方法としては real-time MRI (rtMRI) が近年用いられ始めている。rtMRI は他の調音運動測定器と比べて声道の全体像が高い空間解像度で把握できる反面、音声および画像に強いノイズが重畳され、また時間分解能が比較的低いという欠点がある。先行研究ではノイズ処理されたデータに対して 23.2 frame/sec の時間分解能で推定を行っているが、音声合成ではより高い時間分解能が求められる場合が多い。そこで本研究では、3DCNN により時空間を考慮した特徴抽出を行った上で、1DTransposedCNN (1DTCNN) を用いて時間分解能を上げることによりメルケプストラムの推定精度向上を目指す。実験の結果、表 1 に示すように提案手法が先行手法と比べて高精度な音声合成を実現した。このことから、3DCNN と 1DTCNN の有効性が確認できた。

Table1 メルケプストラムの推定誤差

	メルケプストラム歪み [dB]
先行手法(2DCNN+LSTM)	9.09
2DCNN+1DTCNN+LSTM	6.16
提案手法(3DCNN+LSTM)	5.85

### 1-8-16

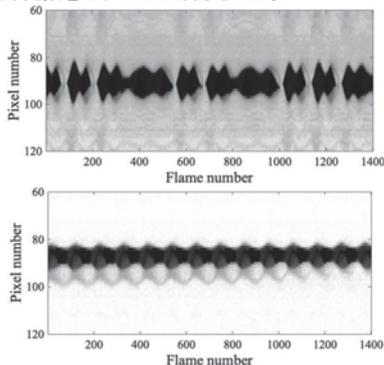
#### 1-8-16 生体パラメータを考慮に入れた仮声帯物理モデルの構築

Construction of a physical model of the ventricular folds that imitate human physiology

☆松本拓磨(立命館大), 金谷麻由佳(立命館大), △長谷川寛人(立命館大)

△松島大輔(立命館大), 石村憲意(立命館大), 徳田功(立命館大)

- ◆声帯および仮声帯の振動を模擬する物理モデルを構築した。
- ◆声帯には MRI モデル、仮声帯には人の仮声帯の形状を考慮した仮声帯物理モデルを作成し使用した。
- ◆仮声帯間の距離を 0mm から 3mm まで 1mm ずつ変化させ、各条件における振動時の声門下圧、音声、高速度動画を計測した。
- ◆仮声帯間の距離が 1mm 以下になると、仮声帯は振動を始め、非周期的な複雑な発振を示すことが確認された。



◆Fig.1 Kymograms observed when the ventricular fold-ventricular fold distance is set to 0 mm (top) and 2 mm (bottom).

### 1-8-15

#### 1-8-15 非定常流条件における声門内圧力分布の計測

Experimental measurement of intraglottal pressure distribution under unsteady flow

☆本田拓人, △井上卓哉, △佐藤玄基, 金谷麻由佳, 徳田功(立命館大), △Anne Bouvet, △Annie Van Hirtum, △Xavier Pelorson (LEGI, UMR CNRS 5519)

- ◆声帯振動時、声帯表面に作用する声門内圧は声帯形状の変化に伴って変化する。この声門内圧の変化が、声帯振動の持続にとって重要であることが知られている。
- ◆本研究では、M5 型声帯レプリカモデルを用いて、声門形状が逆の字型(Divergent 10°)の場合における声門内圧を測定した。
- ◆2Hz で安定した正弦波状で周期振動する非定常流と不規則に乱れる非定常流条件で声門内圧力を計測し、圧力分布と剥離位置の変動について考察した。

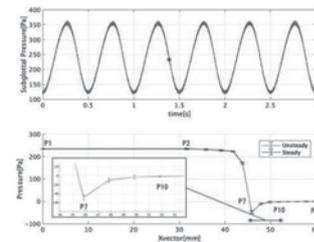


Fig 1 Time series of subglottal pressure (upper panel) and the corresponding pressure distribution (lower panel). Unsteady flow measurement (red) is compared with steady-flow measurement (black)

### 1-9-1

#### 1-9-1 音叉の振動継続時間における個体差の評価

Evaluation of individual differences of vibration duration time among tuning forks

☆野水響太, 大塚翔, 中川誠司(千葉大)

- ◆音叉には、発振周波数の正確さに加えて、発音の十分な持続が求められる。また、音叉を叩いた直後には、基音成分に加えて倍音成分が発生するが、倍音成分は速やかに減衰することが望ましい。
- ◆しかしながら、音叉の製作現場では専ら基音の周波数のチューニングのみがなされており、各成分の振動継続時間の評価は行われていない。
- ◆本研究では、音叉の振動継続時間における個体差と保持強度が振動の持続に及ぼす影響の調査を目的として、同種(同型番)の複数の音叉から発生する音の評価を行った。音叉の保持強度を変化させながら各音叉で計測を繰り返し、基音と倍音の振動継続時間の比較を行った。
- ◆基音および倍音の振動継続時間には有意な個体差が認められた。また、保持強度の変化に応じた振動継続時間の変動が観察された。その変動の様子にも個体間の差異が認められた (Fig. 1)。
- ◆音叉の微小な形状の差異や材料の不均質性が、振動継続時間における有意な個体差を生じている可能性が示唆される。

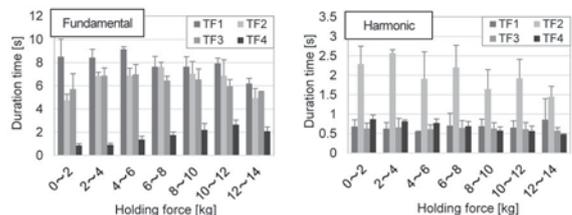


Fig. 1 Duration time of fundamental and harmonic at each holding force range.

### 1-9-2

#### 1-9-2 基底共有型非負値行列因子分解に基づく楽器音の共通・固有成分の分析

Analysis of common and individual components in musical instruments based on basis-shared nonnegative matrix factorization

☆香西海斗, 北村大地 (香川高専)

目的: 複数の楽器音信号の音響的違いや芸術的違いを, 客観的かつ定量的に表現・評価する方法の確立

方法: NMFを用いて複数楽器音信号間の「共通する音響特徴量」及び「固有の音響特徴量」を同時に推定・抽出するアルゴリズムを提案

結果: 2種類のピアノ音を題材として実験し, 両ピアノ音の共通・固有成分を抽出及び考察

今後: 楽器音信号の音色変換, 識別, 検索等への応用方法の検討

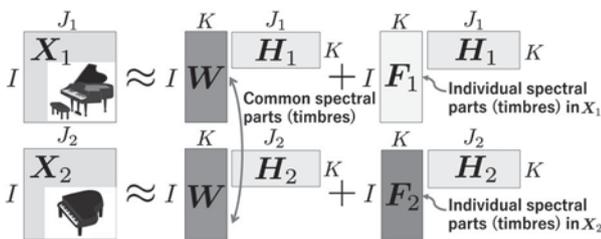


Fig. 1: Decomposition model in proposed basis-shared NMF.

### 1-9-4

#### 1-9-4 非負値行列因子分解を利用した楽曲からの高精度楽器音分離に向けた基礎的検討

\* A fundamental study on high accuracy separation of instruments sounds in music by using non-negative matrix factor decomposition

☆前田梨江, 星野博之 (愛知工業大学)

◆目的: 複数の楽器音で構成された楽曲から個々の楽器音を抽出することを目的とした。非負値行列因子分解(NMF)を使用して楽器音が音源分離によりどのように音質劣化するかを調べた。

◆方法: 2曲の音楽データの約11秒間ずつを切り取り、NMFによって楽器音ごとに分割し合成する。合成音について楽器音ごとの原音に対する明瞭度劣化と発生するノイズを主観評価する実験を行った。実験では基底数、手法、楽器数について比較検討する。手法は KL-divergence 距離尺度条件 (KL) とスパース制約条件 (SP) の2通り、楽器数は2楽器と3楽器の2通りの条件で実験を行った。実験1では基底数は10、20、30の3通り、実験2では深層学習を利用して、基底数を40、50、60とした場合を比較する。

◆結果: Table 1に実験1の基底数10、20、30の結果を示す。実験1、2より2楽器構成の場合、基底数については50が最も良く、ピアノに関してはKL条件で評価値0、コントラバスはSP条件のとき評価値0.5となり最も良かった。

Table 1 Results of evaluation of intelligibility degradation by number of divisions

	ピアノ		コントラバス		
	原音との比較平均	ノイズの平均	原音との比較平均	ノイズの平均	
KL	基底数10	0.3	0.2	1.1	0.8
	基底数20	0.1	0.4	1.8	1.2
	基底数30	0.1	0.4	0.9	1.1
SP	基底数10	1.4	0.9	1.4	1.4
	基底数20	0.5	0.2	0.7	0.8
	基底数30	0.4	0.6	0.5	1.5

### 1-9-3

#### 1-9-3 多チャンネル音響上での分解音声を用いた音脈モアレ表現における群化特性

Grouping characteristics of sound stream moiré using decomposed voice on multi-channel acoustics

○河野有美, △窪田勇樹, 小坂直敏 (東京電機大)

◆マルチチャンネル上での音響空間エフェクトとして音脈モアレ表現の検討を行った。

◆正弦波モデルを用いて音声を分解し, 3メートル四方に配置した4チャンネルのスピーカに各音源を割り当てて再生した。

◆無響室と防音室で実験を行い, 9箇所, 4方向から受聴した時の聴こえた音源の個数と, スピーカの位置を集計した。

◆その結果, 受聴方向や受聴位置だけでなく, 音源の種類によって, 聴こえ方に違いがあった。

◆また, 音響空間の中央位置から受聴した時, どの受聴方向でも1つに群化していなかった。

◆倍音ごとに集計した各受聴方向の平均可聴率をFig.1に示す。

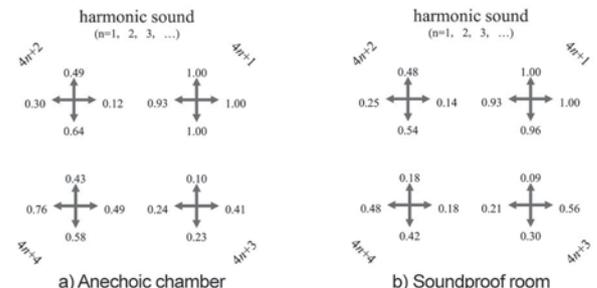


Figure 1 Average number of audible responses for each decomposed voice in each listening direction

### 1-9-5

#### 1-9-5 深層学習を用いた音楽の和音進行の推定

Estimation of chord progression of music using LSTM

○陳 偉博, 小坂直敏 (東京電機大)

◆入力した音楽譜面から, 自動的にその音楽の和声進行を推定するモデルの検討を行った。

◆GTTMのグルーピング構造規則を用いて, 譜面を自動的にセグメンテーションをするアルゴリズムを実装した。

◆クロマベクトルの概念を参照し, それ音名に適用した疑似クロマベクトルを用いて譜面情報を表現した。

◆DNNとLSTMの2段階構造の和音推定システムを提案し, ベートーヴェンの32ソナタ曲を訓練データとしてモデルに学習させた。

◆同作曲家のバリエーションを入力し, セグメンテーションと和音推定の結果を検討した。



図1. 24バリエーションのNo.15変奏曲の推定実験結果  
赤い文字の結果は誤りと判断された結果

### 1-9-6

#### 1-9-6 マルチタスク学習を用いた U-Net に基づく楽曲音源分離に関する調査

An Investigation on Music Source Separation Based on U-Net with Multi-task Learning

☆大竹徹郎(名大), 関翔悟(NTT), 戸田智基(名大)

- ◆楽曲音源分離とは、多様な楽器音が混合した楽曲から特定の楽器音を推定し取り出す技術である。歌声分離においてマルチタスク学習を導入し、歌声のF0推定とU-Netに基づく歌声分離を同時に学習する手法が提案され、高い分離性能が確認されている。
- ◆本稿では楽曲内の様々な音源に対し、マルチタスク学習を用いたU-Netに基づく楽曲音源分離を適用し、その有効性を調査する。また、学習方法の異なる様々なバリエーションのモデルを提案し、マルチタスク学習手順の影響についても調査する。
- ◆実験的評価により歌声、ベース、ピアノにおいてマルチタスク学習を用いることで最も良い分離精度を達成することを確認した。

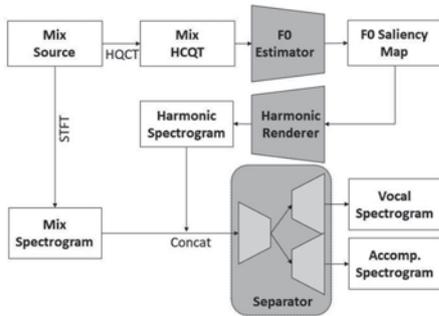


Fig.1: A multi-task learning architecture based on F0 estimation and vocal separation

### 1-9-8

#### 1-9-8 印象の変化に基づいた音楽音響信号の自動分割に対する検討

Automatic separation of musical audio signal based on exciting time.

○橋本まどか(国音大・院), △高倉怜花, 三浦雅展(国音大)

- ◆近年、音楽配信サービスの発展に伴い、楽曲推薦に関わる楽曲の印象の推定に関する基礎技術が盛んに研究されている。その多くは、楽曲に対応する印象は楽曲全体を通して1つであり、楽曲の印象の流動的な変化には着目していない。
- ◆本報告では、楽曲の流動的な印象として、“楽曲の盛り上がり”に着目する。そして、流動的な印象を推定する前段階として、盛り上がりと聴取者が判定する区間の音響パラメータによる推定を試みる。
- ◆このとき、楽曲の盛り上がり区間は同じ小節の間では変わらないという前提を置き、テンポ情報を考慮した音響パラメータを用いた。
- ◆音楽の専門教育を受けた3名の被験者による、盛り上がりと判定する区間のラベリングの結果、有意水準1[%]で相関が認められることが確認できた。
- ◆被験者毎の盛り上がり区間の判定を目的変数、音響パラメータを説明変数とした機械学習を行なった結果、平均79.1[%]で推定が可能であることが確認できた。
- ◆さらに、推定に用いた説明変数に対する重みより、楽曲の盛り上がりの判定については、音の大きさのような盛り上がり区間の判定においては、音の大きさといった音楽の単純な要素を重要視する被験者、スペクトルの形状といった音楽の複雑な要素を重要視する被験者が混在していることが示唆された。

### 1-9-7

#### 1-9-7 音響機器のLSTMベースモデリングにおけるミニバッチ学習のデータ構造に関する一考察

A Study on the Data Structure in Mini-Batch Training of LSTM Machine Learning for Musical Instruments.

☆松永悠斗, ○青木直史, △土橋宜典(北大), △小嶋徹也(東京高専)

我々は機械学習を用いた音響機器(主に歪み回路)のデジタルモデリングについて研究している。

機械学習を扱う上で学習データは重要な要素であり、我々は入力データと教師データのサンプル数に注目し、入力データのサンプル数については検討が済んでいる。

本稿では、教師データのサンプル数について検討する。

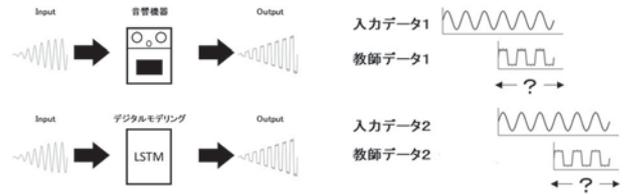


図1. 本研究のコンセプト

図2. 本稿の課題(教師データのサイズを検討)

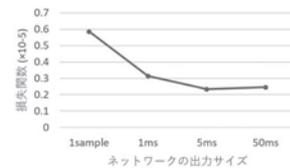


図3. ネットワークの出力サイズ(教師データのサイズ)の違いによるモデリング精度の違い

### 1-9-9

#### 1-9-9 インターネットを利用した楽曲のコード推定の高精度化に関する一検討

A Study on Improving the Accuracy of Musical Chord Estimation by Using the Internet

☆買手 健太, 青木 直史, 土橋 宜典(北大)

昨今、インターネットを介したウェブサービスとして、人手によって作成されたコード譜が無料で公開されるようになってきている。本研究では、こうしたウェブ集合知を利用することで、機械による自動コード推定の誤りを訂正し、コード推定を高精度化する試みについて検討している。本稿では、これまでの予備検討によって得られた結果について報告する。

		1				2				3				4			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1				Em7	D				CM7								D
				Em7	Dsus4				C69								Dsus4
2	Em7				D				C								D
	Em7				D				C69								D
3	Em				D				C								D
	Em7				Dsus4				C69								Dsus4
4	Em				D				C								D
	Em7				Dsus4				C69								D
5	C				D				Em7								Bm7
	C69				Dsus4				Em7								Bm7

図1 Chordify による自動コード推定の結果と楽器.me が公開している人手によって作成されたコード譜の比較(紅蓮華 / LiSA)

### 1-9-10

#### 1-9-10 音声収録における 頭部運動の Kinect による測定 Measurement of head movements using Kinect in voice recording

☆今枝文彦, 片岡優太, 矢田部浩平, 及川靖広 (早大理工)

- ◆ Azure Kinect を用いた発声者の身体運動の計測プログラム
- ◆ タイムコードを介して音声収録システムとの同期収録が可能
- ◆ 発声者の位置と向きの変化を本システムで実際に計測
- ◆ 計測した頭部位置や回転角度から実際の発声動作をアバターで再現

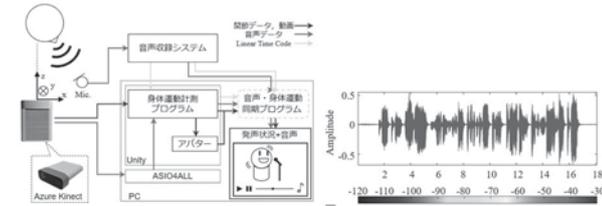


Figure 1 Synchronized recording system for voice and body movement.

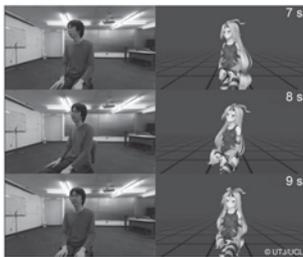


Figure 2 Reproduction of vocal behavior.

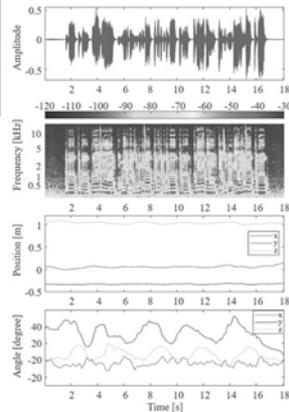


Figure 3 Recording data of voice and head movement.

### 1-9-12

#### 1-9-12 (招待講演) 楽器の音の計測と信号処理

Measurement and Signal Processing of Musical Instruments Sound

○及川靖広 (早大理工)

- ◆ 近年、光学技術やセンシング技術の発展が著しく、それらを用いることでこれまでできなかったことが可能となったり、困難であったことが容易になった。
- ◆ 我々は並列位相シフト干渉法を用いた偏光高速干渉計や透過型 Mixed Reality デバイスを用いた音場情報可視化計測システムを開発してきた。特に前者はマイクロホンを設置しにくい場所での音響計測や楽器近傍や口元の音の計測にも用いてきた。
- ◆ 音響計測や音響信号の収録方法も多様化しており、さらに信号処理の適用により物理現象の理解、音場の微細解析、新しい情報の取得とその利用が可能となる。
- ◆ これまでの研究成果をもとに計測システムとして商品化まで進めることができた。 <https://www.photron.co.jp/solution/20201202.html>
- ◆ 本研究分野に多くの方が興味を持っていただければ嬉しい。

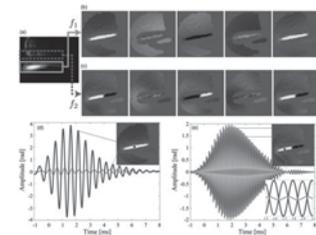


Fig.1 Measurement and Signal Processing of Castanets Sound. Reference source is Kenji Ishikawa, Kohei Yatabe, Yasuhiro Oikawa, "Seeing the sound of castanets: Acoustic resonances between shells captured by high-speed optical visualization with 1-mm resolution," J. Acoust. Soc. Am., vol.148, no.5, pp.3171--3180, Nov. 2020.

### 1-9-11

#### 1-9-11 マリンバとシロフォンの音響分析と音響合成 Sound Analysis and Synthesis of Marimba and Xylophone

Sound Analysis and Synthesis of Marimba and Xylophone

○青木 直史 (北大)

シンセサイザによる音作りに見られるように、音響合成のノウハウは暗黙知のテクニックとして伝えられることが多い。こうしたノウハウを形式知としてまとめていくことは、音響学の発展に寄与するところが大きく、本研究では「ほぼあらゆる音をつくる」プロジェクトと銘打って取り組んできた。本発表では、マリンバとシロフォンを具体例として検討した結果について紹介する。

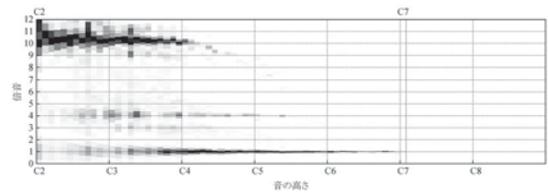


図1 マリンバの倍音構成 (C2~C7)

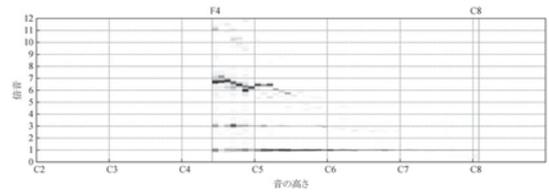


図2 シロフォンの倍音構成 (F4~C8)

### 1-9-13

#### 1-9-13 サウンドエフェクトと電子音楽制作 Sound Effect and Electronic Music Creation

Sound Effect and Electronic Music Creation

○小坂直敏 (東京電機大学)

- ◆ 近年のデジタルサウンドエフェクトで、図1に示す構造的音色に属するエフェクトを中心に紹介する。
- ◆ エフェクトの誕生には技術者側からのみならず、制作側からも提言可能である。
- ◆ その例として擬似エコーと名づけたエフェクトを紹介する。
- ◆ サウンドカラーージュ法の応用として、音源を分解した音色を空間音響上に再生して音脈モワレ現象を起こす空間エフェクトの構想を紹介する。

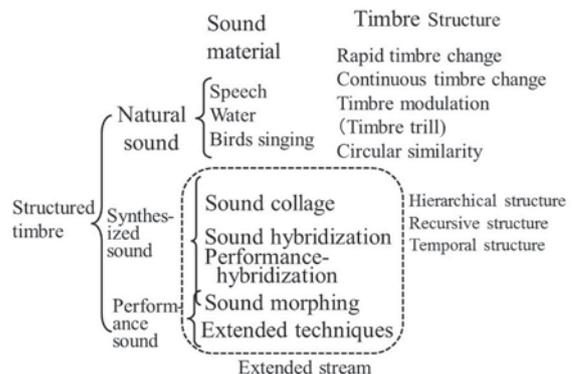


Fig. 1 Classification of Structured Timbre

1-9-14

1-9-14 平面波スピーカーによる立体音響システムの活用

Use of surround system with plane wave speaker

○古山俊一(平成音楽大学)

- ◆平面波による無指向性スピーカーシステム WRAPSOUN NS-8Pro のサウンドとの出会い。
- ◆平面波による無指向性スピーカーシステム WRAPSOUN NS-8Pro は立体音響に応用できるかの実験を行う。
- ◆「富田勲という宇宙」というトリビュートコンサートでの実践的使用例を紹介。サラウンド音源「惑星」より「水星」ダミーヘッドによるバイノーラル再生。
- ◆平面波による無指向性スピーカーシステム WRAPSOUN NS-8Pro を使ったの音場創生。実践例を紹介。サラウンド音源「サン・サース作曲 交響曲 オルガン付き」ダミーヘッドによるバイノーラル再生。

1-10-2

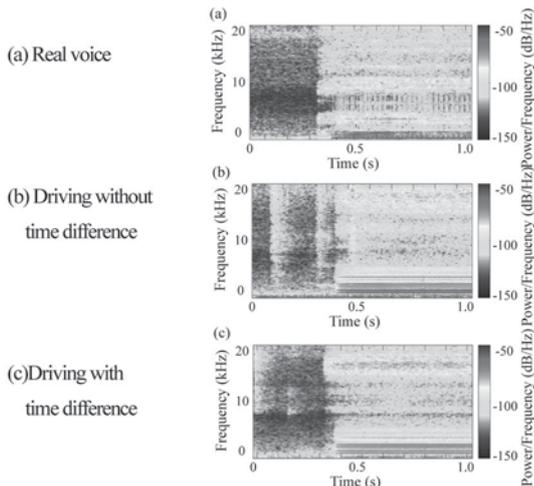
1-10-2 非周期成分を有する駆動信号を用いる電気式人工喉頭

Electrolarynx using driving signal with aperiodic components

☆中村幸寛(筑波大院・シス情工),

若槻尚斗, 水谷孝一, △前田祐佳(筑波大・シス情系)

- ◆電気式人工喉頭の加振部にボイスコイルの振動を直接入力する方式を採用し、健常者の発声に近い駆動信号を検討した。健常者の発声の立ち上がりの部分に特徴があることから、非周期成分の付与を遅らせることで健常者の発声に近づける事を試みる。
- ◆駆動波形の非周期成分をパルス成分より早く入力する事で、健常者の発声の息もれ声による立ち上がりを表現できる可能性が示唆された。



1-10-1

1-10-1 簡易声質維持機能を持つウェアラブルな構音障害支援システムの音声処理装置の検討

Improvement in sound processing device of a wearable voice correction system with a simple voice quality reproduction function

◎藪謙一郎, 伊福部達(東大)

- ◆音声・言語の障害は、発声機能、構音機能、それらの制御機能の何れの障害でも起こり、それぞれ適切な支援策は異なる。例えば、喉頭摘出者は人工喉頭を用いて発声機能を補えば、残された構音機能を使って発話できる。
- ◆筆者らは、構音機能について、リアルタイムに自由度の高い音声を生産できる方式を目指し、指の位置や動きで声を楽器のように声を操作できる音声生成器を開発している。この方式を利用した、ウェアラブルな構音障害の支援システムの試作と、簡易的な声質維持機能を提案し示してきた。
- ◆本稿では、この構音支援システムの試作機で生じていた問題に対処するための、音声処理装置の改良試作について報告する

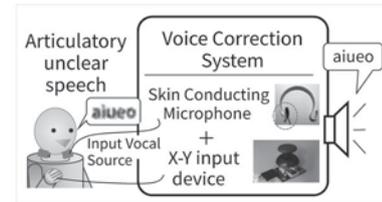


Fig.1: The voice correction system

1-10-3

1-10-3 音響法によるポータブル型耳管機能検査装置の試作

Development of apparatus for the portable eustachian tube function test based on sonotubometry

○朝倉巧, 石井太智(東京理科大), 三輪正人(はりまざクリニック)

- ◆耳管機能不全に起因する耳管開放症などの疾患の診断において、耳管の生理的機能を把握するための手法である T. T. A. G. 法、加圧減圧法、音響法などの手法が提案されている。しかしながら、これらの手法を適用するために必要となる検査装置は大型であり、手軽な計測および診断が困難である。そこで、小型音響デバイスを用いて、音響法によるポータブル型耳管機能検査装置を試作した。本報では、試作器の基本的性能について報告する。
- ◆サンプリング周波数 48 kHz にて透過音の測定を行い、フィルタリング処理によって濾波することで、外耳道内で収録した透過音の相対音圧レベルを求めた。その結果、被検者毎に透過音のレベルは異なるが、耳管開大時にレベルが上昇し、耳管が閉鎖すると再びレベルが下降する様子が確認できた。

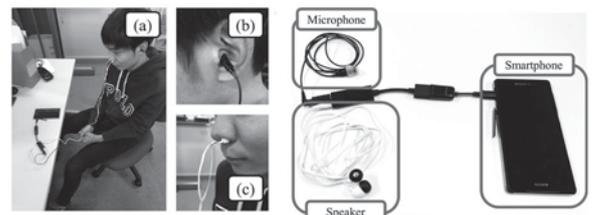


Fig.1: Apparatus of the proposed system.

### 1-10-4

#### 1-10-4 機械学習を用いた 室内環境音の識別に関する検討

Examination for identification about indoor environmental sounds  
using machine learning

☆中屋雅文, 朝倉巧(東京理科大)

- ◆聴覚障がい者を環境音検知の観点から支援することを目的とし、日常生活における数種の室内環境音を、機械学習によって識別させる。
- ◆2種類の住居において日常生活で発生しうる室内環境音15種類をそれぞれ1秒間測定し、これらを教師データとした学習機を用いて室内環境音の識別結果に関して考察した。
- ◆実際に日常生活で聞き取られる環境音は一連の流れの中で発生するものであるため、これを測定した。測定した音声データを1秒間に切り出し問題データとして一連の流れの中での室内環境音の識別結果に関して考察した。

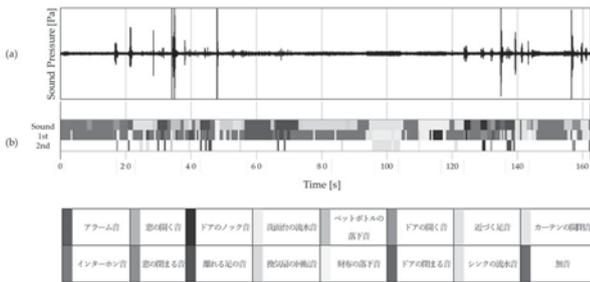


Fig.1: (a) Measurement results of continuous environmental sounds, (b) Classification results of each environmental sound

### 1-10-6

#### 1-10-6 熱動作・熱音響冷却機とリニアモーターを 接続したシステムによる 冷却機温度調整に関する実験

Experimental study on cooler temperature adjustment of a heat driven thermoacoustic cooler by a linear motor

◎千賀麻利子, 長谷川真也(東海大・総技研)

- ◆本研究では、ダブルループ型熱動作・熱音響冷却機にリニアモーターを接続し、冷却機温度を調整するシステムを提案し、実験にて検討を行った。
- ◆実験では熱動作・熱音響冷却機が動作している条件でリニアモーターへ印加する電力  $W_L$  を変更し、その際の冷却機蓄熱器低温側温度  $T_R$  を測定した。
- ◆実験結果を Fig. 1 に示す。横軸に  $W_L$ 、縦軸に  $T_R$  とした。  $W_L = 0$  W はリニアモーターを駆動しないときの結果である。図より  $W_L$  を増加させるごとに  $T_R$  が低下する結果が得られた。  $W_L = 4.9$  W では  $T_R = -18.0^\circ\text{C}$  となり、リニアモーターを駆動しない場合に比べて  $20^\circ\text{C}$  以上低下した。

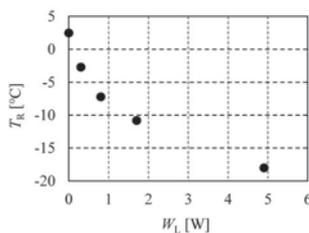


Fig.1 Experimental Results.

### 1-10-5

#### 1-10-5 発達特性のある幼児・児童を対象とした音育活動 -イヤートレーニングの開発を例にして-

Sound Education Activities for Preschoolers and Children with Developmental Characteristics

-Using the example of the development of ear training-

○小松正史(京都精華大学) △華園力(はなぞのクリニック)

- ◆発達が気になる子(発達特性のある幼児・児童)を対象として、音育プログラムを実際に体験した親御さんの声を中心に報告する。
- ◆音楽聴取をはじめ、音を使うことで、子どもの興味を強く引き出すことが可能である。
- ◆22名の回答を得た。年齢層は6~12歳で、男性11名、女性11名であった。
- ◆発達特性のある幼児・児童は、周囲の音環境の状況によって、音の聞き取りが大きく影響する傾向がある。
- ◆音を通して、主体的に身体感覚を育てていくことと、気持ちのよい生活環境を無理のないペースで創り上げることの両方が大切である。

章	トレーニング	効果を感じた割合
第2章-5	紙をやってみよう	62%
第2章-2	新しい足音を出してみよう	55%
第3章-付録	発達が気になる子の環境音楽	50%
第3章-2	楽器の数を数える	48%
第2章-18	気に入った音を探してみよう	39%
第2章-14	音を聴かしてみよう	38%
第3章-10	どんな音が出たかを答える	33%
第3章-5	音を同時に聞き取る	26%
第2章-28	音の違いを考えてみよう	18%
第2章-25	呼吸の音を探してみよう	17%

Fig. The training you experienced and its effects

### 1-10-7

#### 1-10-7 濡れた蓄熱器を用いたダブルループ型 熱動作・熱音響冷却機の臨界温度計算

Numerical study on critical temperature of double-loop type heat driven thermoacoustic cooler with wet regenerator

◎千賀麻利子, 長谷川真也(東海大)

- ◆本報告では Fig. 1 に示すような熱動作・熱音響冷却機の低動作化を目指し、沸点の低い R134a を用いて蓄熱器を濡れた状態とした場合の、原動機蓄熱器高温端面温度に対する冷却機蓄熱器低温端面温度を数値計算によって検証した。また乾いた蓄熱器を用いた場合に対しても同様の計算を行い比較した。
- ◆計算結果を Fig. 2 に示す。同図では冷却機における蓄熱器の低温端面温度  $T_R$  を  $-50 \sim 10^\circ\text{C}$  の範囲で変更し、その際の原動機における蓄熱器の高温端面温度  $T_H$  を数値計算によって求めた。本計算条件では、全計算領域において  $T_H$  は乾いた蓄熱器よりも濡れた蓄熱器で低い結果が得られた。  $T_R$  が  $50^\circ\text{C}$  の場合には  $T_H$  は乾いた蓄熱器で  $243.4^\circ\text{C}$ 、濡れた蓄熱器で  $47.8^\circ\text{C}$  となり、濡れた蓄熱器を用いることで  $T_H$  が  $190^\circ\text{C}$  程度低下する結果が得られた。

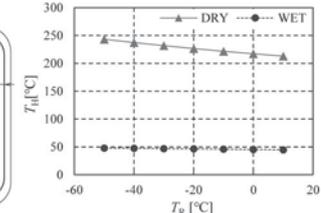
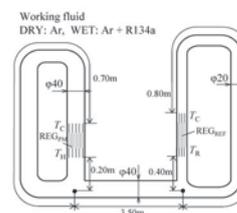


Fig.1 Calculation model. Fig.2 Calculation results (Critical Temperature).

## 1-10-8

### 1-10-8 フルイダインをベースとした濡れた蓄熱器及び吸水芯を有する熱音響機関の臨界温度比の蓄熱器流路径依存性

Regenerator channel diameter dependence of critical temperature ratio of a thermoacoustic engine with wet regenerator and water absorption wick, based on a fluidyne

☆竹村郁哉(東海大院), 千賀麻利子, 長谷川真也(東海大)

- ◆本報告ではフルイダインのように装置内部に液柱を有しながら、吸水芯を用いることで液柱から毛细管現象によって蓄熱器へ液体供給を行い「濡れた蓄熱器」を維持可能な熱音響機関を構築した。構築した熱音響機関を対象に蓄熱器流路径を変更しながらそれぞれの臨界温度比を計測した。また同実験装置を対象に、吸水芯を除くことで「乾いた蓄熱器」を用いた場合の臨界温度比を計測し、比較を行った。開発した熱音響機関の有効性を濡れた蓄熱器(Wet reg.)乾いた蓄熱器(Dry reg.)を用いた場合と臨界温度比の比較によって確認した。
- ◆計測結果を Table 1 に示す。蓄熱器流路径 0.67 mm, 1.18 mm のいずれにおいても濡れた蓄熱器の方が乾いた蓄熱器よりも低い臨界温度比であることを確認した。

Table 1 Measurement results of critical temperature ratio

Channel diameter [mm]	$T_H / T_C$	
	Wet reg.	Dry reg.
0.67	1.17 ( $T_H = 70^\circ\text{C}$ )	1.44 ( $T_H = 150^\circ\text{C}$ )
1.18	1.17 ( $T_H = 70^\circ\text{C}$ )	1.36 ( $T_H = 125^\circ\text{C}$ )

## 1-10-10

### 1-10-10 電気-力学-音響の連成による熱音響デバイスの非線形解析

Non-linear analysis of thermoacoustic device by combining circuit, motion and wave equations

○倉井陽太郎, 上田祐樹(東京農工大学)

- ◆パルス管冷凍機は冷却部に可動部がないため、コスト・整備性・寿命の面で優れた冷凍機である。ゆえに、パルス管冷凍機は極低温冷凍機として宇宙や超電導の分野において注目されている。
- ◆本研究では冷凍機の各要素の非線形性を考慮した熱音響解析を行った。今回は共鳴管と蓄熱器のみの系について数値計算し、実機と比較した。
- ◆印加電圧を固定し、電圧の周波数を変化させた際の共鳴管端部の圧力振幅について求めた。Fig.1 に数値計算と実験での測定結果との比較を示す。周波数が変化しても計算結果はおおよそ一致することが確認できた。

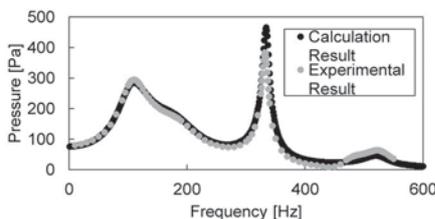


Fig.1 Comparison of Calculation and Experiment

## 1-10-9

### 1-10-9 ジェットポンプの音響特性の測定

Measurements of acoustic characteristic on jet pumps

○小川雄大, 上田祐樹(東京農工大)

- ◆熱音響冷凍機は音波を利用してヒートポンプ効果により冷熱を得る装置である。構成部品が少ない単純な構造をしているなど多くの利点がある。しかし、入力する音波の振幅が大きくなることで生じる音響流が熱音響冷凍機の性能を低下させるため、導波管にジェットポンプを接続することで時間平均圧力差を発生させ、性能低下を防止する手法がある。ただジェットポンプは万能ではなく、管の形状による損失であるマイナーロスが生じる。そこで本研究では時間平均圧力差を維持しつつ、マイナーロスをなるべく小さくできるジェットポンプの形状を調査した。
- ◆形状を変えて作成したジェットポンプに対して、差圧係数と損失係数を算出し、レイノルズ数との関係を調査した。レイノルズ数がある程度大きくなると、差圧係数と損失係数は小さくなりながら漸近し、ほぼ一定の値をとることが確認できた。

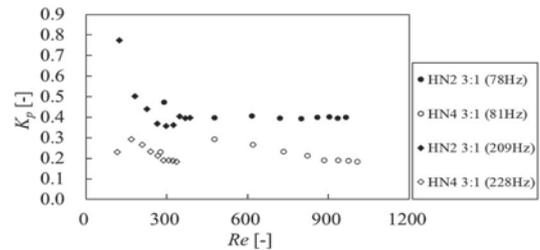


Fig. 1 Differential pressure coefficients as a function of Reynolds number

## 1-10-11

### 1-10-11 熱音響現象の有限振幅時の予測における伝熱面積モデルの検討

Consideration of heat transfer area model in prediction of finite amplitude of thermoacoustic phenomenon

☆大橋謙一郎, 上田祐樹(東京農工大)

- ◆熱音響デバイスは工場排熱や太陽熱などの未利用熱の有効利用手段として注目されている。熱音響デバイスは熱音響理論により解析がされている。熱音響理論では作動流体の温度を求めることができるが、実際には熱交換器の温度を変化させ、エンジンを駆動している。本研究ではこれまでに提唱されている熱交換器の伝熱面積や熱伝達率モデルを統合し、有限振幅時の共振温度予測に適用することで熱交換器温度を求め、実機と比較した。
- ◆作動流体の熱伝導率と熱交換器の水力直径の比に係数をかけた値を熱伝達率とし、それを適用して数値計算で熱交換器の温度を求めた。Fig.1 に実験で測定した熱交換器の温度との比較を示す。係数は 1.5 もしくは 2 とし、実験では熱交換機の内周と外周の温度を測定した。係数が 1.5 の場合は中心と外周の平均値に一致し、2 の場合は中心部の温度に一致した。

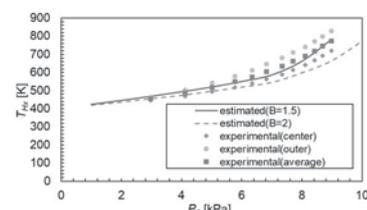


Fig. 1 Comparison of heat transfer coefficient

### 1-11-1

#### 1-11-1 猫用低騒音被毛回収手段

How to collect of fur with low-soundlevel for cats

○馬場智子, △新井達也, 藤原奨(三菱電機)

- ◆猫の清潔性維持を目的に日常的にブラッシングを行うが、このときに問題となる猫の被毛の空中飛散の抑制と飼い主のストレス改善を両立する猫専用の製品を開発した。
- ◆猫は人間よりも聴覚が優れ、外来音に対して敏感なため、猫に触れて使う電気製品を開発するにあたり、猫の音に対する快/不快を定量的に確認するための聴感実験を行った。
- ◆聴感実験の結果、周波数帯域と音圧の組合せが、猫の音に対する快/不快を決定づけることが示唆された。
- ◆Table 1に聴感実験における試験音に対する猫の反応結果を、Fig.1に試作機を使って猫の被毛を回収している様子を示す。

Table 1 Reaction of cats against the test sound

試験音の概要		猫の反応結果
1/1oct. 帯域音	250 Hz	反応なし
	500 Hz	不快ではない
	1 kHz	不快
	2 kHz	
	4 kHz	
8 kHz		
単一周波数音	10 kHz	不快
	15 kHz	
	17 kHz	
	20 kHz	
	25 kHz	



Fig.1: The picture of collecting cat fur by using the prototype

### 1-11-3

#### 1-11-3 気筒間ばらつきの音圧変動に基づくディーゼルエンジン燃焼音の快適化

Comfortable sound design of diesel engine combustion noise based on sound pressure fluctuation of variation among cylinders

☆戸田勇介(中央大院), 戸井武司(中央大)

ディーゼルエンジンは気筒内圧の上昇がガソリンエンジンより急激に変化するため、燃焼音が目立ち、各気筒における音響伝達関数や燃料噴射タイミング等の違いによるばらつきが音質に影響を与える。

そこで本研究では、気筒間ばらつきに着目し、各気筒における爆発順序を変え、燃焼音の印象変化を把握する。Table 1に3つのクラスターのエンジン回転の0.5次成分の変調度(DM0.5ord)、1次成分の変調度(DM1ord)と主観評価の相関分析で異なる結果を示す。DM0.5ordで正の相関、DM1ordで負の相関と回転次数による傾向を把握した。また、周波数帯によって相関の大きさが異なる。次に、周波数帯毎の音圧変動と感性の影響を把握するために、振幅変調波を用いた主観評価を実施した。Fig. 1に振幅変調波の音質評価の結果を示す。周波数帯によって変調度による影響が異なることを示し、3000-6000 Hzで音圧変動が感性に与える影響を大きくすることを確認した。さらに、音圧変動を表す特徴量と燃焼音の主観評価から重回帰分析を用いて、燃焼音の快適感推定モデルを構築し、音圧変動に着目した燃焼音の快適感の指標を示した。

Table 1 Correlation analysis

Metrics	Cluster		
	1	2	3
DM0.5ord (Full Band)	0.54	0.52	0.38
DM1ord (Full Band)	-0.50	-0.44	-0.15
DM0.5ord (0-1500 Hz)	0.49	0.46	0.34
DM1ord (0-1500 Hz)	-0.21	-0.22	-0.07
DM0.5ord (1500-3000 Hz)	0.75	0.75	0.45
DM1ord (1500-3000 Hz)	-0.10	-0.18	-0.04
DM0.5ord (3000-6000 Hz)	0.55	0.56	0.36
DM1ord (3000-6000 Hz)	-0.65	-0.64	-0.29

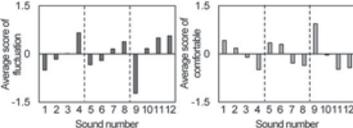


Fig.1 Evaluation result (Amplitude modulation)

### 1-11-2

#### 1-11-2 加速意思に応じた加速感を得られるエンジン音の創生

Creation of engine sound that gives a feeling of acceleration according to the intention to accelerate

☆清水周作(中央大院), 戸井武司(中央大)

本研究では走行条件として運転者の加速意思に着目し、意思に見合った加速感が得られるようなエンジン音の創生を目指す。加速意思はドライビングシミュレータの走行タスクで区別を行うものとし、前方車を追い抜く場合を大(Aw\_L)、気にせず走る場合を小(Aw\_S)とする。

まず、3種類のエンジン回転数上昇率[大(Res\_L)、中(Res\_M)、小(Res\_S)]を用い、各加速意思での印象変化を把握した。

Fig. 1に示すとおり、加速意思によらずエンジン回転数上昇率が高いほど快適感が得られやすく、加速感に関しては各加速意思に対し適切なエンジン回転数上昇率があることが分かった。

次に、3種類のエンジンの音色[無加工(Def)、低次成分強調(Low)、高次成分強調(High)]を用い、各加速意思での印象変化を把握した。

Fig. 2に示すとおり、加速意思によらず低次成分を強調することで加速感が得られやすいが、快適感が下がる傾向がみられた。さらに、加速意思が小さい場合では快適感の変動が大きくなることが分かった。

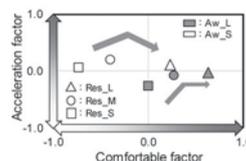


Fig. 1 Evaluation result of changing engine speed increase rate

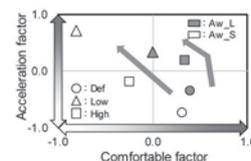


Fig. 2 Evaluation result when changing the tone of the engine sound

### 1-11-5

#### 1-11-5 オノマトペ表現による ロータリースイッチ操作感覚の評価

Evaluation of rotary switch operation feeling by onomatopoeia expression

☆宮入徹(中央大院 / 都産技研), 勢川真以(中央大),

△下村尚登, △白坂剛(アルプスアルパイン), 戸井武司(中央大)

- ◆ロータリースイッチ操作時の感触および操作音の印象に対して、自由回答によるオノマトペ評価を行った。
- ◆実験結果に対してテキストマイニングによる分析を行い、操作時の印象から連想されやすいオノマトペの傾向について調査した。
- ◆感触に関しては、回転時のクリックトルクの大きさによって用いられる語の傾向が異なった (Fig.1)。
- ◆操作音に関しては、音量および音質によって用いられる語の傾向が異なった。

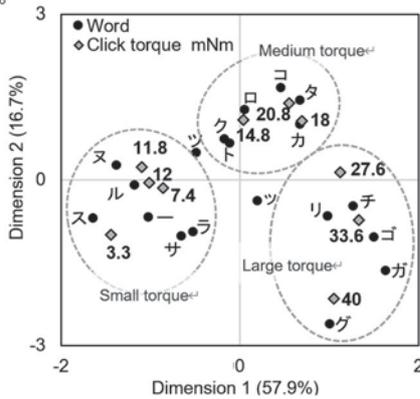


Fig.1: Correspondence analysis results (haptics)

### 1-11-7

#### 1-11-7 使用状況に応じたキーボードの 操作感覚の創生

Creation of keyboard's operation feeling according to playing condition

☆藤倉圭(中央大学院), 戸井武司(中央大学)

メカニカルキーボードはスイッチ軸ごとに様々な操作感覚が存在するが、操作目的に適した青茶赤軸は明確になっていない。そこで本研究では、各スイッチ軸における感触や操作音に関わる操作感覚を把握し、操作目的別に適した操作感覚の構築を目指す。

Fig. 1に各軸の操作目的による主観評価を示す。連打時には感触にタクトイルを有する茶軸の快適感が高く、長押し時にはタクトイルレスな感触を有する赤軸の快適感が高いことがわかる。これより各スイッチ軸で操作目的によって明瞭感、快適感に変化することがわかった。

次に、Fig. 2に青軸の操作音の1000Hz付近のピーク音圧を±20dB加工した音源を用いた操作目的による主観評価を示す。連打時には、明瞭感の高いクリック感を有するグループのオリジナル音や+20dB操作音の快適感が高まり、長押し時には、明瞭感の低いグループの-20dB操作音の快適感が向上した。本実験から連打時には明瞭感の高い操作音や感触が好まれ、長押し時には明瞭感の低い操作音や感触が好まれることがわかった。

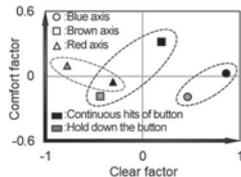


Fig. 1 Evaluation results of operational objective by SD method

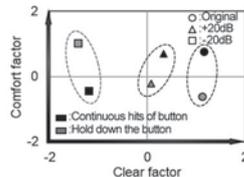


Fig. 2 Evaluation results of operation sound by SD method

### 1-11-6

#### 1-11-6 ロータリースイッチの物理特性を考慮した 操作音の印象評価

Evaluation of operation sound considering physical characteristics of rotary switch

☆勢川真以(中央大), 宮入徹(中央大院/都産技研),

△下村尚登, △白坂剛(アルプスアルパイン), 戸井武司(中央大)

- ◆ロータリースイッチは、その構造上、スイッチの切替動作時にクリック音と摺動音が生じ、聴感上の印象にはそれぞれの音が寄与している。
- ◆そこで本研究では、クリック音および摺動音の物理特性を個別に操作することで、クリック音と摺動音がどのように印象に寄与しているかを実験的に検討した。
- ◆結果から、操作音の音量やシャープネスの変化が、快適性、重厚性、および明瞭性の印象に影響を与えることがわかった。
- ◆また、クリック音と摺動音の音量のバランスを適切に調整することで、快適性を向上し得る可能性を示した。

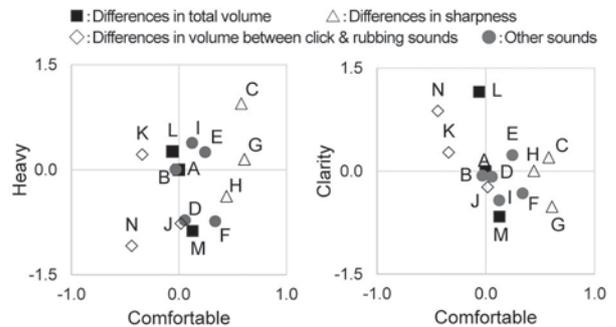


Fig. 1: Evaluation results of operation sound by Participants

### 1-11-8

#### 1-11-8 自動または手動変速時における 音と振動を考慮した操作感の向上

Improvement of operational feeling by sound and vibration in automatic or manual gear shifting mode

☆野島惠理子(中央大院), 戸井武司(中央大)

本研究では、主観評価を用いて変速時のエンジン音や変速ショックと変速音のバランスおよび変速音の周波数特性による印象変化を検討する。

まず、エンジン音と変速音のバランスの主観評価を行い、次に変速音の周波数特性の違いによる主観評価を行った。これらの結果より、変速音を付加することで自動変速(AT)よりもパドルシフト(PAD)の方が操作感が向上することがわかった。また、快適感において、エンジン音と変速音のバランスおよび変速音の周波数特性が影響していることが示唆された。

最後に、変速音と変速ショックのバランスの主観評価を行い、因子分析を行った。Fig. 1に変速ショック中(m)の操作因子×快適因子の散布図を、Fig. 2に変速音と変速ショックの寄与率を示す。上記の評価と同様に変速音を付加することで操作感が向上することがわかり、変速音が操作感に影響していることが示唆された。加えて、変速ショックが快適感に影響を与えていることも示唆された。

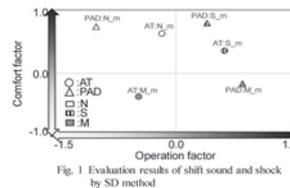


Fig. 1 Evaluation results of shift sound and shock by SD method

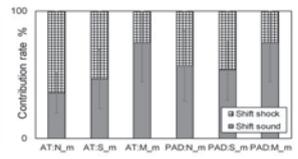


Fig. 2 Contribution rate of shift shock and shift sound

### 1-11-9

#### 1-11-9 セカンダリアクティビティの作業負荷に応じた警告音の選定

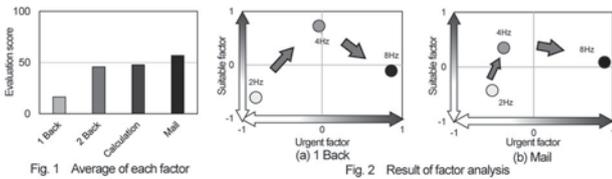
The selection of warning sounds corresponding to the degree of workload of the secondary activity

☆五十部健太(中央大院), 戸井武司(中央大)

自動運転中は運転動作とは別の作業, すなわちセカンダリアクティビティ(SA)を行うことが可能になる。本研究では, SA の作業負荷の大きさに応じた適切な警告音の選定を行う。

まず, 1 Back 課題, 2 Back 課題, 計算課題, メール課題の4種類において作業単体での負荷の大きさを NASA-TLX を用いて評価を行う。Fig. 1 にそれぞれの作業の6項目における評価結果の平均値を示す。作業の種類によって作業負荷に差異が生じることが確認でき, SA 評価に用いられる妥当性を確認した。

次に, 自動運転中の SA として, 作業負荷が大きく異なる 1 Back 課題とメール課題を選定し, 警告音の有用性を主観評価で行う。Fig. 2 にそれぞれの課題における緊急因子×適合因子の散布図を示す。吹鳴周期が短くなるにつれ緊急感が上昇した。また, 負荷の小さい 1 Back 課題では, 吹鳴周期が 4Hz と 8Hz の適合感に顕著な差異が生じることを確認した。以上より, 手動運転への権限移譲時に SA による作業負荷に応じた警告音を提示する必要があることが推察される。



### 1-11-11

#### 1-11-11 報知音の協和感および音列パターンが警告感に及ぼす影響

Effects of urgency by harmony and several patterns of auditory alarms

☆阿部理花子, 大塚翔, 中川誠司(千葉大)

- ◆災害無線, 緊急地震速報やエリアメールなどの災害情報発信システムの利用が進んでいる。音による災害速報では, アナウンスの口調と同等に冒頭に呈示される“報知音”の役割が大きい。
- ◆警告効果の大きい報知音の設計を目指して, 構成音の協和感および音列パターンが警告感に及ぼす影響を対比較にて調査した。
- ◆まず, 構成音の周波数比を変化させることで協和感を変化させ, 警告感との関係を調べた (Table 1)。各被験者においては, 警告感協和感に応じて変化したが, 被験者間のばらつきが大きかった。
- ◆次に音列の時間パターンと警告感の関係を調査した。刺激長は長い方が警告感は大いだが, 持続時間 200 ms で飽和すること, および繰り返し周波数が高い方がより警告感が増加することが示唆された。基本的にはエネルギーが大きく, 音響イベント (オンセット) の数が多い方がより警告感が強くなると考えられる。

Table 1 Triads used in the experiment 1

sound	周波数 [Hz]			周波数比
A	420	630	840	2 : 3 : 4
B	420	525	630	4 : 5 : 6
C	420	490	560	6 : 7 : 8
D	420	472.5	525	8 : 9 : 10
E	420	462	504	10 : 11 : 12
F	420	455	490	12 : 13 : 14
G	420	450	480	14 : 15 : 16

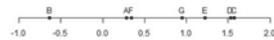


Fig. 1 Scale value for the urgency in the experiment 1

### 1-11-10

#### 1-11-10 前後を表す音のデザインに関する一考察

A Study of Sound Design for Representing Front and Back

吉野華織, 和氣早苗 (同志社女子大)

【目的】サイン音としての利用を想定し, 言語や定位を利用しない音そのもののデザインで前後を表現できるか検討を行う。

【実験概要】「ピッチの高低」「ピッチ変化(上行, 下行)」「和音(長和音, 短和音)」を変化させた実験音を用いて, 前後認識に関する比較評価と絶対評価を実施した。

【結果】◆単音のピッチ高低と前後認識: ピッチ差が大きいほど「高ピッチを前」と認識する可能性が高く, また提示順序の影響も受けにくいことが示唆された (図1)。

◆ピッチ変化による前後認識: ピッチ変化の中心ピッチが同じ場合は「上行が前」の傾向が見られたが, それよりもむしろ中心ピッチの高低による影響(「高ピッチが前」)が強く作用する。また別の観点として「安定感の前」という傾向も示唆される (図2)。

◆和音と前後の認識: 絶対評価では, 長和音, 短和音いずれについても, 前という回答が後ろという回答を上回った。

◆高ピッチ, 上行, 安定感が「前」という傾向が示唆された。これらの組合せでより確実な前後の表現ができるか, 今後研究を深めたい。

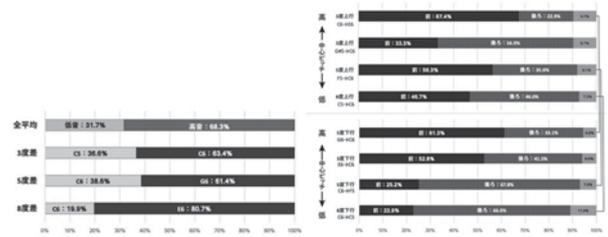


図1. 単音ピッチ高低と「前と回答」した割合 図2. 上行/下行, 中心ピッチと前後認識

### 1-11-12

#### 1-11-12 機械学習による音源の特徴量選択に基づいた快音化モデルの構築

Construction of comfortable sound model based on selection of sound source features by machine learning

☆後藤悠太(中央大院), 戸井武司(中央大)

工業分野では音質評価を基に人の感性と音源物理量を結びつける快音化モデルの構築が行われている。従来, 熟練者による特徴量決定や評価が行われてきたが, 特徴量の選択ミスや取りこぼしの可能性がある。

本研究では機械学習を用いて, 従来手法では選択できない特徴量を抽出し, 予測精度の高い重回帰モデルを作成し, 従来モデルとの比較検討を行った。従来モデルでは, 熟練者が聞いて快適感に関係すると考えられる特徴量からモデル構築を行っている。提案手法では, 機械学習のランダムフォレストで算出される変数重要度を用いて, 特徴量を抽出し, 重回帰分析に使用した。従来モデル(C-model)と提案モデル(P-model)の修正済決定係数 $R_f^2$ を Table 1 に示す。 $R_f^2$ はモデルの推定精度を示し, 0.8 以上で信頼性が高い。また, 検証データに対する精度評価指標を用いた予測精度を Table 2 に示す。P-model は, 誤差を示す指標が低く, 二乗平均平方根誤差(RMSE)と平均絶対誤差(MAE)の比から適切なモデルと言える。また t 検定よりモデル間の平均予測誤差に有意差を確認した。C-model と P-model の差は有意なものであり, 精度が高いことがわかった。

Table 1 Adjusted coefficient of determine

Model	C-model	P-model
$R_f^2$	0.767	0.901

Table 2 Evaluation index

Model	C-model	P-model
MAE	0.571	0.386
RMSE	0.644	0.476
Standard Deviation	0.297	0.279
RMSE/MAE	1.127	1.235

### 1-11-13

#### 1-11-13 食欲に対する聴覚と視覚に基づくモデル構築

Model building based on hearing and vision for appetite

☆黒瀬和希(中央大院), 戸井武司(中央大)

近年、遠隔地への情報発信が聴覚、視覚情報により行われている。食欲に着目すると、視聴者の食欲が増す聴覚、視覚情報の提示条件はあまり研究されていない。そこで本研究では、食材(肉)を焼く調理音と調理映像に着目し、音または映像の特徴を変更した場合の被験者の食欲変化を評価実験より把握し、食欲のモデル構築を行う。

Fig. 1に音または映像の単刺激に対する空腹度、Fig. 2に音と映像の複合刺激に対する空腹度を示す。Fig. 1より、周波数帯域は、No.3の1000~2000 Hzが好まれ、再生速度はNo.5の0.50倍が好まれる。色合いは、G10が好まれる。周波数帯域と再生速度を比べると、再生速度の方が影響度大きいことが分かる。視覚情報よりも聴覚情報における空腹度が高く、聴覚情報が有利であるといえる。Fig. 2より、聴覚情報の同一で視覚情報の違いによる空腹度の差分平均は2.0、視覚情報の同一で聴覚情報の違いによる空腹度の差分平均は3.6で、複合刺激下においても視覚情報に対し、聴覚情報が有利であるといえる。

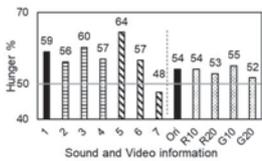


Fig. 1 Hunger for each single stimulus

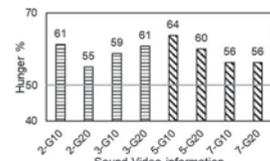


Fig. 2 Hunger for each combined stimulus

### 1-11-15

#### 1-11-15 せせらぎの主観的好ましさに寄与する音響パラメータの検討

Effects of acoustic parameters of murmuring of a stream on subjective preference

☆高橋貴行, 大塚翔, 中川誠司(千葉大)

- ◆シャープネスがせせらぎの聴感印象に与える影響を一对比較によって検討した。
- ◆市販の音源DVDに収録されたせせらぎの再生速度を1/3~3倍に変化させることで、段階的にシャープネスを変化させた7種類のせせらぎを作成した(Table 1)。Scheffeの一对比較法を用いて、シャープネスがせせらぎの聴感印象に与える影響を検討した。
- ◆多くの先行研究と異なり、シャープネスが一定値より小さい場合には、却って好ましさが低下した。
- ◆シャープネスが小さい刺激音では、せせらぎにふさわしいシャープネスや流速の範囲を逸脱してしまったために、不自然さが増してしまった可能性が考えられる。

Table 1: Stimulation sound information

Stimulation sound	Reproduction speed [double]
A	1/3
B	1/2
C	2/3
D	1
E	3/2
F	2
G	3

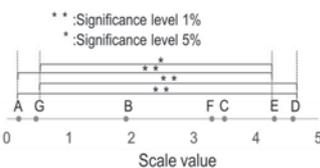


Fig. 1 Averaged scale value of the preference for the stimuli.

### 1-11-14

#### 1-11-14 深層学習と非負値行列因子分解を用いた自然環境音からの快適音創生

Creation of comfortable sounds from various natural sounds by using deep learning and non-negative matrix factorization

○星野博之, △江崎諒平(愛知工業大学)

◆目的 運転中などの人が活動する各種作業環境において、人の状態を適切かつ快適に保つための音刺激を自然環境音から探索、創成することを目的としている。今回は、多数の自然環境音のオクターブバンドスペクトログラムと快適さ主観値を深層学習によりモデル化し、非負値行列因子分解(NMF)を用いて生成した自然環境音の部分音の快適さを推定することにより快適音を探索した。

◆方法 自然環境音 250音を学習データとし、その中の主観値が高い25音の1kHz, 2kHz, 4kHz帯域抽出音(計75音)をテストデータとして、畳み込みニューラルネットワークを用いて深層学習を行った。次に、NMFを用いて上記25音を、いろいろな分離・生成数の条件で部分音に分解し、深層学習モデルにより快適さを推定した。

◆結果 Fig. 1に上記25音の元音の出力値と25音の部分音の最大出力値を示す。結果として、今回の手法により自然環境音から快適音を生成できることが示された。

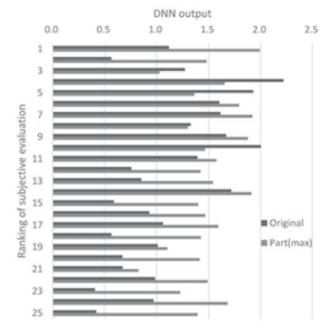


Fig.1 Natural sounds comfort estimates by DNN

### 1-11-16

#### 1-11-16 図書館を題材とした「自分事」化プロジェクトの実践その2

音環境からみた大学図書館における「アノニマス性」に関する検討

A case report on project "struggle as his/her own issues" in KAIT library part 2: "Anonymity" in university library from soundscape perspective.

○小橋佑生(神奈川工科大), 三浦貴大(産総研), 土田義郎(金沢工業大),

△田岡壮平, △松本一教, 田中哲雄, 上田麻理(神奈川工科大)

前報[1]より、静かさと賑やかさという物理的に対極にある音環境が求められているが、その本質は「アノニマス性」という共通性であることが分かった。

そこで本稿では、図書館のアノニマス性の構成要因を明らかにするために質問票調査を行った。さらに、図書館で存在する環境音や自己発音及び、他者の発音の心理的・物理的な影響を探るために、キーボードを用いた文章入力タスク実験を実施した。



Figure: Example of application screen and application configuration

### 1-11-17

#### 1-11-17 残響デザインによる空間的印象変化に関する研究

Study on spatial impression changes by reverberation design

☆前崎壮真(中央大院), 戸井武司(中央大)

空間的印象としての拡がり感に注目し、周波数特性および時間特性を操作する残響デザインによる主観量の印象変化を把握し、得られた主観量を目的変数に用いた拡がり感推定モデルを構築する。

まず、SD法による評価から空間的、美的、明るさ、柔らかさの4因子を抽出し、空間的因子と残響時間の変化に類似した傾向を確認した。次に、残響時間の周波数特性を把握し、2種の手法を用いて体感残響時間を補整し、拡がり感向上への効果を把握した。最後に、基準音に加え、体感残響補整を施した上で周波数および時間特性を操作した計13種の音源を用いたシェッフェの対比較法(浦の変法)による評価を実施し、これらに対する拡がり感の平均嗜好度を得た。これより、重回帰分析を用いた拡がり感推定モデルを構築した。Table 1に作成した2種のモデルを、Fig. 1に○で示す評価値と各推定モデルの比較を示す。自由度修正済み決定係数の値から1000Hz帯域の残響時間を用いたModel2の予測精度が高く、これによる拡がり感向上への効果が検証できた。

Table 1 Estimate model

Metrics	Model 1	Model 2
RT (1000Hz Band)	-	+0.43
RT (All Band)	+0.35	-
Articulation Index	-	-0.85
Sharpness	-	-10.8
Impulsiveness	+175.8	-
Constant term	-15.7	+109.2

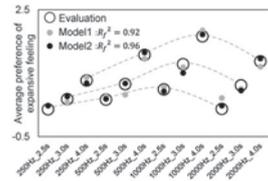


Fig. 1 Evaluation value and 3 types of estimation values

### 1-2P-2

#### 1-2P-2 口唇口蓋裂者の音声認識のためのデータ拡張方式の検討

An investigation of data augmentation method for speech recognition of cleft lip and cleft palate

☆富士原健斗, 高島遼一(神戸大), 杉山千尋, 田中信頼, 野原幹司, 野崎一徳(大阪大), 滝口哲也(神戸大)

- ◆口唇口蓋裂者は口唇や口蓋に割れ目が生じているため、発話が開鼻声と呼ばれる聞き取りにくい音声になる。
- ◆口唇口蓋裂者音声の学習データを大量に収録することは困難であり、健常者に比べてラベルが不正確になる可能性も高い。
- ◆本研究では、データ拡張を用いた学習データの増量や、人手によるラベルに依存しない自己教師あり学習を行う。
- ◆健常者を対象にした従来手法に加え、開鼻声の特性を考慮した新たなデータ拡張「周波数伸縮」を行った結果、特定話者音素認識実験における誤認識率が低減することが明らかになった。

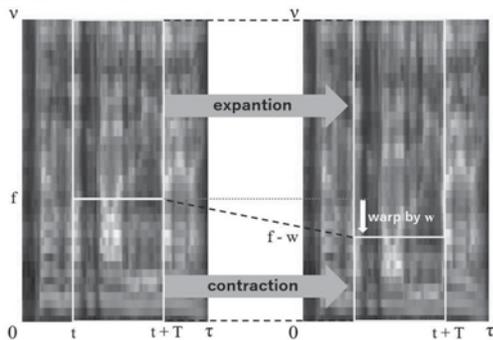


Fig. 1: Overview of our proposed method.

### 1-2P-1

#### 1-2P-1 マルチモーダル音声認識におけるLocal attentionを用いた音声画像統合方式

Audio-image integration using local attention for audio-visual speech recognition

☆角田遼太(神戸大), 相原龍(三菱電機)

高島遼一, 滝口哲也(神戸大), 本山信明(三菱電機)

- ◆AV Alignは音響特徴量をクエリとして画像特徴量にAttentionをかけるCross-modal attention機構により音声と画像の特徴量を統合する手法である。
- ◆AV AlignはCross-modal attentionを計算する際に、全ての時刻の画像フレームに対して重み付けを行っているが、雑音環境下において正確な重みの推定が困難な可能性がある。
- ◆本研究では、ある時刻の音声に関する画像のフレームは全体の一部であるという考えに基づき、重みの計算を一部の画像フレームに限定するLocal attentionを使用することを提案する。
- ◆Local attentionを用いることで、特に大きな雑音が存在する環境下では従来手法よりも優れた結果を示すことができた。

Table 1 Character Error Rate (CER) in background noise

Model	Window size	CER[%]		
		clean	10 dB	0 dB
audio only	-	24.8	35.7	52.8
baseline(global)	$\infty$	22.0	31.3	46.6
local	5	21.8	30.3	44.1
	11	21.2	29.2	42.7
	21	21.1	29.5	43.2
	31	21.7	29.2	43.8

### 1-2P-3

#### 1-2P-3 Efficient Channel Adaptation of ASR by DNN-based Data Augmentation using Re-recorded Paired Data with Automatic Alignment Correction

自動音声認識の録音チャンネルの適応のためのDNNに基づくデータ拡張と再収録音声の自動アライメント修正

☆Raufun Nahar, Atsuhiko Kai (Grad. School of Science and Tech., Shizuoka Univ.)

- ◆This paper focuses on data augmentation approach by simulated features using feature transformation for adaptation of acoustic model
- ◆Training feature transformation model using re-recorded audio as target requires precisely to be paired with original audio data. Temporal misalignment in re-recorded speech causes fault in training
  - Linear misalignment (delay) caused by clock drift between playback device and recording device
  - Variable misalignment caused by transmission delay or packet loss in telephone channel
- ◆Correcting alignment at starting point of re-recorded data improves performance of channel adapted acoustic model remarkably

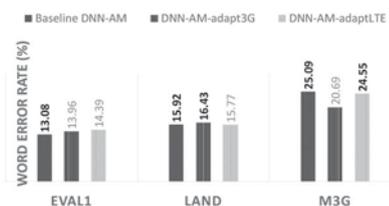


Fig. 1: Evaluation (WER%) of ASR adaptation for 3G and LTE mobile channels with data augmentation

1-2P-4

講演取消

1-2P-5

講演取消

1-2P-6

1-2P-6 テキスト情報を利用した深層学習に基づく  
音声感情認識

Speech Emotion Recognition Using Linguistic Features  
based on Deep Learning

☆永瀬 亮太郎, 福森 隆寛, 山下 洋一(立命館大)

【概要】

音声とテキストの情報を考慮した深層学習に基づく感情認識の提案

【目的】

- (1) 認識率の更なる向上
- (2) 音声以外の情報との統合的な処理技術の研究

【提案手法】

事前に学習した音声情報とテキスト情報についての感情認識モデルを用いて、各認識器から得られる出力に対し early-fusion と late-fusion を適応する

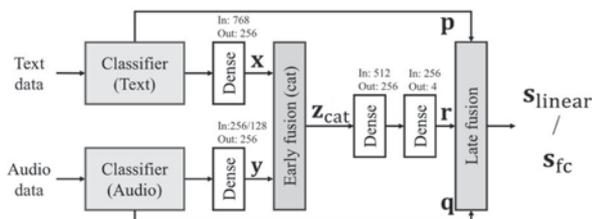


Fig.1: Network Architecture of Speech Emotion Recognition Using Linguistic Features

【結果】

- 音声・テキスト情報単体の認識器よりも全体的に認識率が向上
- late-fusion を適用することで、最大20%高い認識率を獲得

1-2P-7

PPG を用いた  
日本語 TTS システムの構築

Building a Japanese TTS system by using PPG

○孫慶華, 永松健司(日立)

- ◆近年、音声認識、対話生成、機械翻訳などの技術と組み合わせた音声合成サービスの需要が急速に高まっている。特に他社製品や応用シーンの違いを差別化するために、複数の異なる声質や話し方を使い分けられることのできる、音声合成製品が要求されている。そのため、合成ベンダにとっては「声質」と「話し方」(以降まとめて「話者性」と呼ぶ)の異なる合成音声を大量に提供すること、すなわち話者性の「量産化」が市場シェア拡大の鍵となる。
- ◆通常、新規話者の音声合成を実現しようとする、ターゲット話者の音声を収録した後に、音声に対するラベリング作業が必要となる。特に収録音声に対する書き起こし作業は、発話内容を忠実にテキストで表記する必要があるため、ここに大きなコストがかかり、話者性の量産化を阻害する要因となっている。
- ◆そこで本研究では、音声に含まれている言語情報(特に発音情報)を忠実に表現できる音素事後確率(PPG: Phonetic PosteriorGram)を用いた音声合成手法を提案し、ラベリング作業を必要としない音声合成システムの構築手法を実現した。
- ◆音声合成実験では、3000人以上の大規模音声コーパスを用いて音声合成システムを構築し、高い合成音質を実現したうえで、新規音源による音声合成システムの自動構築を実現した。
- ◆今後は、音声に含まれる韻律情報のEmbeddingを用いて、合成音声の自然性向上を目指す。

### 1-2P-8

#### 韻律を考慮した end-to-end 方式に基づく 自発音声合成

Prosody-aware end-to-end spontaneous speech synthesis.

☆西野広直, 森大毅 (宇都宮大)

##### ◆ 目的

自発音声のバラ言語情報を制御するためには、韻律パラメータである基本周波数(F0)の時間変化パターンを適切にモデリングすることが必要である。

##### ◆ 手法

本研究では end-to-end 音声合成の1つである Tacotron 2 において F0 とスペクトルのマルチタスク学習を提案する (Fig. 1)。

##### ◆ 結果

韻律を考慮した Tacotron 2 (Tacotron-F0) と従来 Tacotron 2 の F0 歪み, 自然音声の再現性を評価した。

- Tacotron-F0 の F0 歪みは従来法のものより有意に減少した。
- Tacotron-F0 の選好率が従来法を上回った。

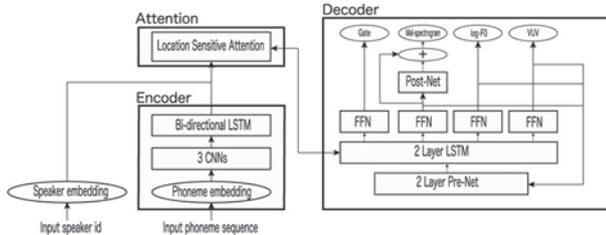


Fig.1: Tacotron-F0 architecture.

### 1-2P-10

#### 1-2P-10 母音発音と歌唱速度の変化を考慮した アカペラオペラ歌声合成

Singing voice synthesis for a cappella opera considering variations of vowels and singing speed.

☆片平健太(神戸大), 足立優司(メック株式会社),

田井清登(メック株式会社), 高島遼一(神戸大), 滝口哲也(神戸大)

- ◆ アカペラオペラ歌唱に現れる特徴を考慮したモデルを従来の歌声合成システムに加える。
- ◆ オペラ歌唱において、顎を大きく開いて発声することより、母音の発音が通常のものから変化する現象がみられる。
- ◆ オペラ歌唱での母音の発音のケブストラムの分布を主成分分析及び k-means クラスタリングによって分析し、その結果を歌声合成の音響特徴量推定モデルの学習・生成に用いる
- ◆ アカペラ歌唱では歌唱速度が歌唱者に依存し、その歌唱速度は楽譜で指示される一定のテンポから頻繁に変化する。
- ◆ 歌唱速度の変化を楽譜の各セグメントの経過時間の変化と捉え、自然なテンポ系列を推定するモデルを歌声合成システムに追加する。

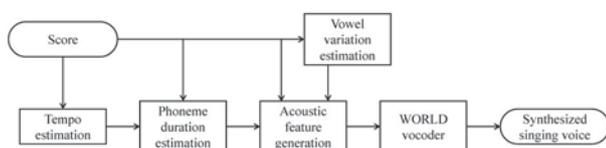


Fig. 1 Singing voice synthesis system for a cappella opera.

### 1-2P-9

#### 1-2P-9 歌唱表現を付与できる Bidirectional-LSTM を用いた 歌声合成方式の検討

Study of singing voice synthesis generate song expression using Bidirectional-LSTM

◎金子隼人, 原直, 阿部匡伸(岡山大・HS 統合科学研)

- ◆ 本稿では、楽譜情報と音素情報を入力とする歌唱表現を付与する歌声合成方式を提案する。
- ◆ 歌唱表現を再現するためには、歌唱表現付きの歌声データが必要となる。しかし歌唱表現付きの歌声データを大量に集めるのは困難である
- ◆ 提案方式では、比較的少量な歌声データとそれに比べて少量の歌唱表現データから学習することを想定している
- ◆ 歌唱表現のない一人話者の約 50 分のデータセットと歌唱表現のある複数話者の約 10 分のデータセットから音響特徴量を分析し、これを正解として学習をおこなった。
- ◆ 複数話者のデータで学習することによる音質への影響を調査するために、複数話者のモデルと一人話者のモデルで音高と声質に関して客観評価実験をおこなった。
- ◆ 歌唱表現のうちピブラートに着目し、ピブラートを付与する歌声合成方式について検討をおこなった。
- ◆ 提案方式によって合成された歌声における基本周波数の時間変化がピブラートの変化パターンに近づいていることを確認した。

### 1-2P-11

#### 1-2P-11 深層距離学習に基づく特徴量を用いたL2発話比較に関する実験的検討

An experimental study on L2 utterance comparison based on deep metric learning

◎安藤慎太郎, 齋藤大輔, 峯松信明 (東大・工)

- ◆ Posteriorgram-DTW: 語学教育システムにおける発話類似度計算技術としても使用される
  - ◇ Posteriorgram (音素クラス事後確率ベクトル系列) 言語依存 → 従来手法だと言語数の音響モデルが必要であり、多言語対応システム構築の課題となる
- ◆ 本研究: 深層距離学習モデルから生成された特徴量を Posteriorgramの代わりに利用することを検討
- ◆ Triplet lossによる音素クラス識別siamese networkを構築
  - ◇ Posteriorgram: 音響モデルで音素クラスを (絶対的に) 認識するよう学習
  - ◇ 音素クラス識別特徴量: 各入力サンプル間の音素クラス的一致/不一致 (相対的な関係) のみを使用
  - ◇ 音素ラベルの「絶対的使用→相対的使用」により、言語依存の「音素」への依存性の低下を期待
- ◆ タスク: 英語・日本語のL2発話比較
  - ◇ DTWスコア↔主観評価等の目的スコア の相関で評価
- ◆ 学習データ: 多言語コーパス (タスク言語を含まない)
- ◆ 異なる話者による2発話比較のタスクでは、多言語音素クラス識別モデルは言語依存モデルの最高性能には及ばないながらも、英・日の両タスクで最も安定した性能を示した

### 1-2P-12

#### 1-2P-12 ビブラート歌唱音声における スペクトル変動消失音声の自然性の評価

Evaluation of naturalness of singing voice without spectral fluctuation in vibrato singing.

◎鈴木千文, 坂野秀樹, 旭 健作(名城大), 森勢将雅(明治大)

- ◆本研究ではスペクトル変動が一定であるビブラート合成音(スペクトル変動消失音声)の作成方法の検討を行った
- ◆歌声データ: ビブラート音声(目標歌手: Gackt)
- ◆スペクトル変動消失音声の作成
  - ビブラート区間の基本周波数の各フレームのスペクトルの値で、ビブラート開始以降のフレームの値を置き換える
- ◆メルケプストラム歪みの算出(Fig.1)
  - 基準フレームをビブラート開始区間のフレームとし、ビブラート中間区間のフレームとの歪みを算出(Fig.1(左))
  - 基準フレームをビブラート終了区間のフレームとし、ビブラート開始区間のフレームとの歪みを算出(Fig.1(右))
  - Fig.1(右)の方が、Fig.1(左)に比べてメルケプストラム歪みの値が全体的に大きく、基準フレーム間における差も大きい
  - ビブラートの終了区間では、スペクトルの形状が大きく変化していることが考えられる

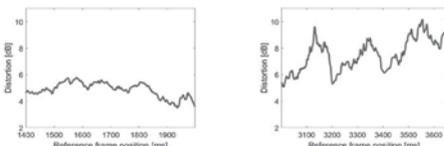


Fig.1 Mel-cepstrum distortion.

### 1-2Q-2

#### 1-2Q-2 スペクトルとケプストラムを併用した 深層学習に基づく悲鳴検知の検討

A study of scream detection based on deep learning with spectrum and cepstrum

◎福森 隆寛 (立命館大)

- ◆研究目的
  - 雑音環境における悲鳴検知に適した音声特徴量の獲得
- ◆研究概要
  - 従来の悲鳴検知では、音声認識などで用いられていた特徴量(メルスペクトログラムやメル周波数ケプストラム係数など)を転用していた。本稿では、Fig. 1のように異なる空間から音声特徴量を個別に獲得する機構(Convolutional neural network: CNN)を導入し、音声の特徴を多角的に捉えながら、悲鳴を検知する手法を提案した。
- ◆評価実験
  - 雑音環境における評価実験を実施し、スペクトログラムとケプストログラム(ケプストラムを時系列方向に並べた特徴量)を入力特徴量とするCNNを用いることで雑音にロバストな悲鳴検知を実現した。

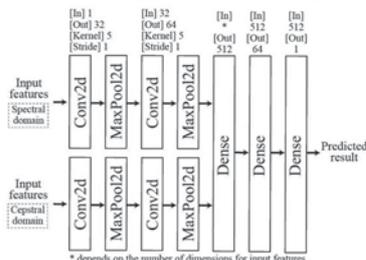


Fig.1: Proposed CNN for scream detection.

### 1-2Q-1

#### 1-2Q-1 Attention モデル Teacher-Forcing を用いた 長時間音声とテキストの自動アライメント

Automatic Alignment of Long-form Speech and Text using Teacher-Forcing for Attention-based model

◎木田祐介, 小松達也, 戸上真人 (LINE)

- ◆End-to-End 音声認識の学習への利用を想定した音声とテキストの新しいアライメント方法を提案する。
- ◆提案法は、二つの異なる Transformer モデルを利用する。はじめに、forward Transformer がテキストを前から順に走査し、発話の末尾に対応するテキストの終端を推定する。次に、backward Transformer が、推定した終端からテキストを逆順に走査して発話の先頭に対応するテキストの始端を推定する。最後に、推定した区間のテキスト系列に対する backward Transformer の事後確率を用いてアライメントの正しさを判定する。
- ◆CSJ を利用した評価実験の結果、提案法はいずれの評価セットに対しても **96%を超える高いアライメント精度**を示した。

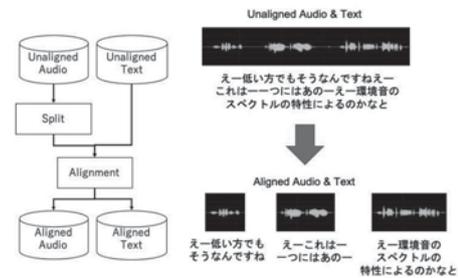


Fig.1: Problem setting of this paper. Our problem is an utterance-based audio-text alignment.

### 1-2Q-3

#### 1-2Q-3 複数エージェントを用いた音声対話システムにおける差別化およびエージェント間対話の効果

EFFECTS OF ROLE ASSIGNMENTS AND INTER-AGENT CONVERSATIONS IN MULTI-AGENT SPOKEN DIALOGUE SYSTEM

◎数崎祐貴, 李晃伸(名工大)

- ◆複数エージェントを用いた音声対話システムを調査し、エージェント間の差別化と対話が用いられることを見出した。
- ◆複数エージェント間の差別化と対話を強化した案内対話システムを実装し、2要素が真に効果的か検証を行った。
- ◆実験の結果、差別化および対話のどちらか一方を取り入れた対話システムでは単体のシステムの方が良い結果となったが、両方を取り入れることにより、単体以上の効果を得られることが示された。

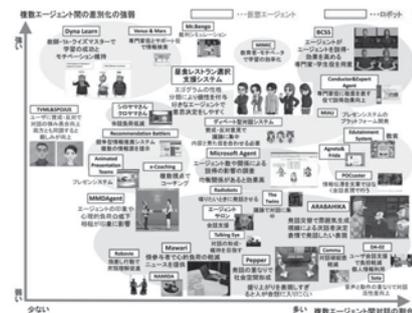


Fig.1: 複数エージェントを利用した対話システム群 (黄色は情報提供, 赤は説得, 緑は負担低下を狙う)

### 1-2Q-4

#### 1-2Q-4 自発音声に基づく合成音声で対話するシステムがユーザに与える影響の調査

How does a spontaneously-speaking dialog system affect users?

☆飯塚喬久, 森大毅, 西野広直 (宇都宮大)

- ◆従来の音声対話システムでは読み上げ音声のコーパスから合成音声のモデルを作っていた。それで良いのか?
- ◆我々は、一般的な合成音声と比べ、自発音声に基づく合成音声で対話するシステムの方が、より社会的な存在と認識されるという仮説を立てた。
- ◆自発音声コーパスを元にした合成音声で話すシステムと読み上げ音声コーパスを元にした合成音声で話すシステムの2種類と人との対話を評価した。
- ◆「どれほど人間同士の会話に近かったと思いますか」の質問項目に対して読み上げ音声コーパスを元に合成した場合(平均3.05)に比べ、自発音声コーパスを元に合成した場合(平均3.58)の方が高い評価を得た。



### 1-2Q-6

#### 1-2Q-6 音高情報条件つき変分自己符号化器を用いた歌唱 F0 パターン生成

Fundamental Frequency Contour Generation of Singing Voice Using Score-conditional Variational Autoencoder

◎関翔梧 (NTT), 多賀遥香, 武田一哉, 戸田智基 (名大)

- ◆喉頭摘出者のための歌唱支援システムにおいて、より自然で多様な歌唱表現の実現を目指して、これまでに変分自己符号化器 (VAE) を用いた歌唱 F0 パターン生成法が提案されている。
- ◆VAE を利用することで、人手で制御した場合と同等の歌唱 F0 パターン生成が可能となる一方で、歌唱 F0 パターンが持つ揺らぎ成分を適切に表現することは困難であり、システム利用時には音高情報とともに入力する雑音成分に対して、手動による大きさの調整が必要となる。
- ◆本稿では、音高情報を補助情報として利用し雑音成分を区別して扱う条件つき VAE (CVAE) を用いた歌唱 F0 パターン生成法を提案する。
- ◆実験的評価より、雑音成分を調整することなく従来法と同等の性能が達成可能であることが確認された。

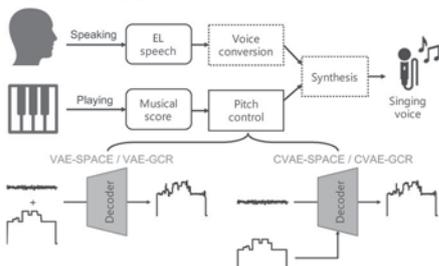


Fig.1: Overview of singing aid system using VAE-based methods (conventional) and CVAE-based methods (proposed).

### 1-2Q-5

#### 1-2Q-5 音響イベント検出と位置推定における転移学習の効果の検証

Investigation on the effect of transfer learning in sound event localization and detection

☆陳軼夫, 山田武志, 牧野昭二 (筑波大)

- ◆音響イベント検出 (SED: sound event detection) と位置推定 (SEL: sound event localization) を同時に行う SELD タスクにおいて、Cao らは転移学習の有効性を示した。しかし、SED と SEL のどちらを転移元とするのか、またネットワークのどの部分を転移対象とするのかについては十分に議論されていなかった。
- ◆本稿では、SELD タスクに対して広く用いられている、CNN (convolutional neural network) と BGRU (bidirectional gated recurrent unit) を組み合わせたモデルを対象とし、転移学習の効果的な適用方法を検証した。
- ◆実験結果を Table 1 に示す。SED タスク、SEL タスク共に、時間構造の分析を担う BGRU を転移することが効果的であり、SED タスクにおいてはスペクトル構造の分析を担う CNN を転移すると性能低下を招くことが明らかとなった。

Table 1 Experimental results for each learning method.

Method	Development dataset				Evaluation Dataset			
	SED task		SEL task		SED task		SEL task	
No transfer	0.66	52.7%	25.6°	59.2%	0.60	55.0%	23.8°	63.2%
SED→SEL (CNN & BGRU)	/	/	15.5°	74.1%	/	/	14.4°	76.3%
SEL→SEL (BGRU)	0.58	55.3%	/	/	0.55	56.3%	/	/
SEL→SEL (CNN)	0.78	35.2%	/	/	0.71	44.2%	/	/

### 1-2Q-7

#### 1-2Q-7 敵対的生成ネットワークを用いた歌声変換の各種検討

Examination of singing voice conversion using generative adversarial networks

☆堀幸央, 佐伯和哉, 小坂哲夫 (山形大)

- ◆敵対的生成ネットワーク (GAN: Generative adversarial networks) を用いた日本語歌唱の歌声変換の検討を行った。
- ◆GAN を用いた声質変換の検討は多数行われているが、歌唱に関する検討は少数に留まるため、各種基礎検討を行った。
- ◆学習に関して、異なる話者間の同じ内容の歌唱を利用する並列学習、および異なる歌唱を利用する非並列学習の2種類が存在する。歌声変換において両者はどの程度差があるか検討を行った。
- ◆変換前後でどの程度目標話者に近づいているか客観評価するために DTW で歌声間の対応づけを行い、ケプストラム距離を算出した。
- ◆Fig.1 より非並列学習では性能が劣化することが分かった。また距離計算より変換音声に大幅に近づくことが分かった (Table 1)。

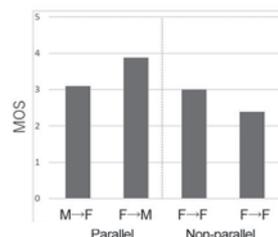


Table 1: Mel-cepstral distance between various voices

	Source ↔ Conversion	Conversion ↔ Target
Female→Male	41.45	20.62
Male→Female	39.43	25.29

Fig 1: Evaluation results of naturalness by MOS (F: female, M: male)

### 1-2Q-8

#### 1-2Q-8 Improving spectral detail and F0 modelling for VAE-based cross-lingual voice conversion with adversarial training

©Tuan Vu HO, Masato Akagi (JAIST)

This paper proposes a non-parallel cross-lingual voice conversion (CLVC) system using Variational Autoencoder (VAE) combined with Star Generative Adversarial Network (StarGAN) that can perform both speaker adaptation and speaker individuality control. In this paper, the StarGAN training scheme is adopted to enhance the performance of the VAE model for CLVC tasks. We also propose an F0 injection method to provide more controllability over the F0. After training on a cross-lingual multi-speaker database, the proposed voice conversion system can adapt an arbitrary source voice to a new voice by interpolating the input speaker embedding. Objective and subjective evaluations confirm the effectiveness of the adversarial training scheme and F0 injection method.

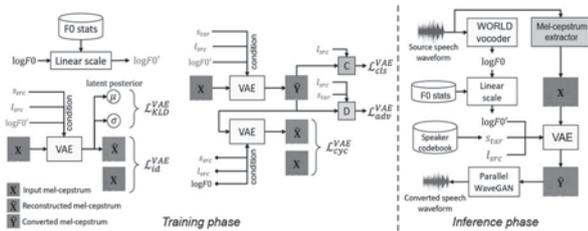


Fig. 1: The overview of processing flow of proposed VAE-StarGAN model. The VAE model acts as a generator of the StarGAN model.

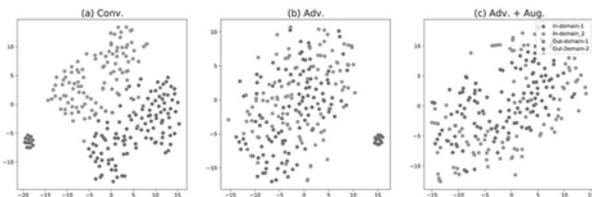
### 1-2Q-10

#### 1-2Q-10 データセットや言語の違いに頑健な話者埋め込み手法とそのテキスト音声合成への応用

Speaker embedding method robust to dataset and language differences and its application to text-to-speech synthesis

☆後藤駿介, 齋藤大輔, 峯松信明 (東大)

- ◆ 話者埋め込みベクトルを補助特徴量として用いる複数話者テキスト音声合成では、話者埋め込み空間は話者の声質のみに依存するべきである。本研究では、データセットや言語が話者埋め込み空間に与える影響を調査するためにバイリンガルデータセットを活用し、言語の違いの影響がデータセットなどの環境の違いの影響よりも軽微であることを示す。そして、データセットの違いに対して頑健な話者埋め込み手法としてドメイン適応とデータ拡張を用い、それぞれの手法の有効性を主観評価実験によって検証する。



学習内外のデータを入力することで得られる話者埋め込みベクトルの t-SNE による可視化

### 1-2Q-9

#### 1-2Q-9 複数話者音声合成の高品質化のためのアクセント変換を用いた韻律正規化の検討

A study on prosodic normalization using accent conversion for high-quality multi-speaker text-to-speech

☆白旗悠真, 齋藤大輔, 峯松信明

- ◆ 問題設定: テキスト音声合成 (TTS) の学習データにおいて、テキストから得られるアクセント情報と実際の読み上げ音声のアクセント位置には齟齬があることがあり、これは生成音声の品質低下に繋がると考えられる
  - 特に複数話者音声合成において、方言などの言語的話し者のばらつきが大きい素人話者の音声を用いた際に顕著
- ◆ 提案法: TTS の学習データの F0 軌跡に対して、テキストから得られるアクセント情報に合わせるようなアクセント変換をかけることで両者の対応関係を改善する
- ◆ 実験結果: 提案法により、TTS 音声においてアクセントの自然性や目的話者との近さが改善できることが示唆された

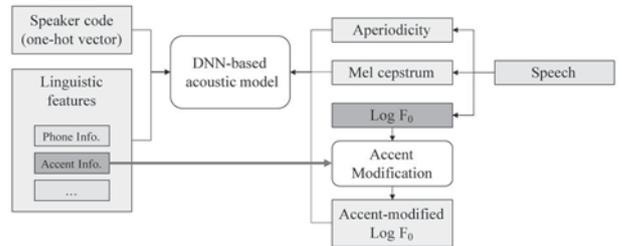


Fig.1: Overview of proposed method. Accent information of Log F0 in training data is modified to match the accent information from text.

### 1-2Q-11

#### 1-2Q-11 深層ガウス過程を用いた sequence-to-sequence 音声合成のモデル構造の評価

Evaluation of Model Structure for Sequence-to-sequence Speech Synthesis Using Deep Gaussian Process

☆中村 泰貴, 郡山 知樹, 猿渡 洋(東大院・情報理工)

- ◆ 以前の報告で、Length Regulator を採用した、音素単位系列から音響特徴量系列へのマッピングが可能な Sequence-to-Sequence DGP 音声合成を提案
- ◆ 本稿では、主観評価により以下を報告
  - 提案法への音素フレーム内位置の付与 & 有効性の評価
  - リカレント機構 + 畳み込み層を持つ DNN モデルと比較し 音声の自然性が向上
  - Self-Attention 機構 + 畳み込み層を持つ DNN モデルと比較し 音声の自然性が向上

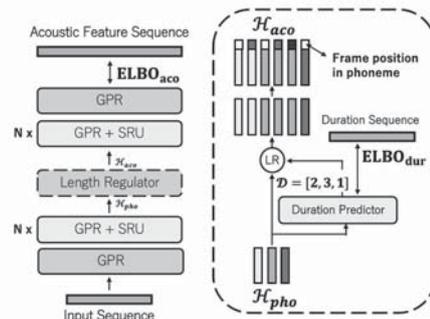


Fig.1: Architecture of Seq2Seq DGP speech synthesis using Length Regulator

### 1-2Q-12

#### 1-2Q-12 Dysarthric Speech Conversion by Learning Disentangled Representations with Non-parallel Data

☆Xunquan Chen<sup>1</sup>, Jinhui Chen<sup>2</sup>, Ryoichi Takashima<sup>1</sup>, Tetsuya Takiguchi<sup>1</sup>,  
(<sup>1</sup>Kobe University, <sup>2</sup>Prefectural University of Hiroshima)

- ◆ In this study, we present a non-parallel Voice conversion (VC) method for dysarthric speech. The proposed framework is shown in Fig.1
- ◆ The model key feature is that, in the training stage, we utilize a phoneme-guided reference encoder (Er) to enforce the phoneme recognition encoder (Ep) to encode the speaker-independent linguistic representation of dysarthric speech. The speaker encoder (Es) accepts the acoustic features as input, and learns the reasonable representations relating to speaker in the embedding space.
- ◆ In the conversion stage, the reference encoder is not used. Using the trained speaker encoder, we can replace the speaker embeddings with target one. The linguistic embedding will be extracted by the trained phoneme recognition encoder. Finally, the decoder (De) generates the converted speech based on the linguistic information in the source dysarthric speech and the speaker information in the target speech.

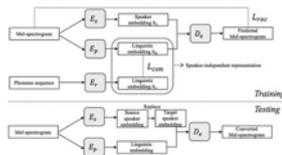


Fig.1: Framework overview.

### 1-4P-2

#### 1-4P-2 聴覚モデルを用いたシャープネスと変動強度の計算モデルの構築

Computational models of sharpness and fluctuation strength using auditory model

☆磯山拓都, 木谷俊介, 鶴木祐史(北陸先端大)

【背景】 これまでに、ガンマトーンフィルタバンク (GTFB) とガンマチャープフィルタバンク (GCFB) をもちいて時間領域計算するラウドネスモデルを提案した。

【目的】 提案したラウドネスモデルを用いてシャープネスと変動強度の計算モデルを提案する。

【やったこと】 提案したラウドネスモデルを用いてシャープネスと変動強度の計算モデルを構築し、シャープネス及び変動強度に関するヒトの知覚データを説明できるか評価した。

【結果】 Table 1, Figure 1 より

Tab. 1 RMSE of each loudness between perceptual data on sharpness and sharpness value from computational models.

Loudness [sone]	2	7	14	28
Fastl & Zwicker	0.260	0.171	0.043	0.102
Aures	0.157	0.124	0.053	0.071
GTFB model	0.105	0.090	0.041	0.042
GCFB model	0.074	0.072	0.044	0.062

- 提案したシャープネスの計算モデルの算出結果は他のモデルの算出結果よりシャープネスに関する知覚データとの二乗平均平方根誤差 (RMSE) の値が小さいことがわかった。

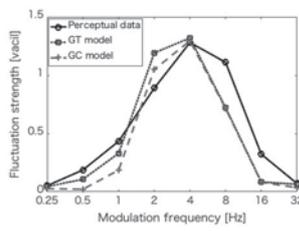


Fig. 1 Fluctuation strength using proposed models for each modulation frequency.

- 提案した変動強度モデルから算出された値は変動強度に関する知覚データと同じ傾向であった。

【結論】 提案モデルは知覚データを説明できることがわかった。

### 1-4P-1

#### 1-4P-1 声のピッチ感の錯覚とスペクトル傾斜の操作による検討

Examination by manipulating spectral gradient in the illusion of voice pitch.

○内田照久 (大学入試センター・研究開発部)

目的 「声の音色」が違っても、「基本周波数が低い」声であっても、「高い声」だと誤って知覚されてしまう錯覚の原因の解明「フォルマント固定時」にも錯覚が生ずるかを検証

方法 WORLD を使用して 声の音色と高さを制御

- (1) スペクトル周波数軸の伸縮 vs 高域強調と低域強調
- (2) 通常  $f_0$  パターンの昇降 vs 平板化  $f_0$  の昇降

結果 (3) 上記を組み合わせた音声の比較対による高さの主観評価

首都圏の国立11大学1年生123名(男:90名・女:32名) スペクトル傾斜による音色の違いは、声の高低判断に影響しかし、その効果は 周波数軸の伸縮条件より小さい可能性

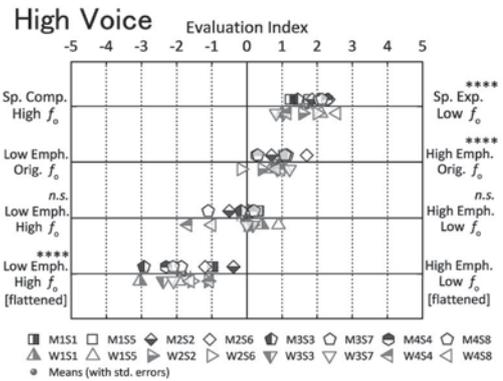


Fig. 1 High voice score for each experimental condition.

### 1-4P-3

#### 1-4P-3 2 純音の同時性知覚における聴覚末梢の寄与

Evidence that perceptual simultaneity is determined at an early stage of auditory pathway

◎岡崎聡, 津崎実(京都市立芸大)

◆聴覚の同時性の範囲(同時性の窓, 時間単位)は周波数次元において「岡崎・一川・津崎(OIT)方程式」に従う。

➢ OIT方程式は両耳同音聴(diotic)条件で得られた。

➢ 一方、両耳異音聴(dichotic)条件では、同時性の範囲はOIT方程式と異なる振る舞いを小さな周波数距離で示す。

◆本研究は両耳同音聴条件と両耳異音聴条件の振る舞いの違いを信号検出理論の観点から分析した。

◆その結果、小さな周波数距離において、

➢ 同時性判断の感度  $d'$  が両条件で異なった。

➢ 同時性判断の判断基準  $C$  は両条件で変化しなかった。

◆これらのことは、同時性知覚が低次の感覚処理により決定されていることを示唆している。

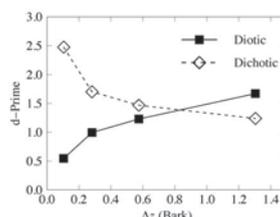


Fig. 1  $d'$  for diotic and dichotic conditions.

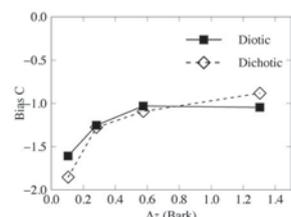


Fig. 2 Bias C for diotic and dichotic conditions.

### 1-4P-4

#### 1-4P-4 聴覚マスキングのための不快音の重畳加算に基づくマスキング音設計

Masker sound design based on overlap addition of discomfort sound for auditory masking

☆小嶋佑弥, 中山雅人, 高橋徹 (阪産大)

- ◆従来の聴覚マスキングに基づく不快感低減手法では、マスキング音として、白色雑音や流水音などが利用されていた。しかし、不快音と特徴が大きく異なる音源を利用してマスキング音を設計し、聴覚マスキングを行うため、不快音とマスキング音が分離して聞こえやすい問題点があった。本稿では不快音を重畳加算し、マスキング音を設計することで不快音との親和性を保持しつつ不快感を低減可能なマスキング音の作成を行う。提案手法では、ヒューマンスピーチライク雑音の作成と同様に、不快音を重畳加算することで、マスキング音の音源を設計する。この重畳加算では、フレーム処理を行い、フレームをランダムに加算する方式と単純に不快音を重畳加算する方式の両方を検討する。
- ◆提案手法の有効性を確認するため、主観評価実験を行った。3種類の不快音(虫の不快音、歯科治療音、ブレーキ音)を対象とし、白色雑音をマスキングとした従来手法(CM)、単純に重畳加算した提案手法(PM w/ SOA)、ランダムなフレームで重畳加算した提案手法(PM w/ ROA)に対して、5段階評価のMOSを算出することで、不快感低減の性能を評価した。Fig.1に虫の不快音, Fig.2に歯科治療音, Fig.3にブレーキ音の実験結果を表す。実験結果より、提案手法は従来手法と比較して、スコアが同等、もしくは高いことから、提案手法の有効性を確認した。

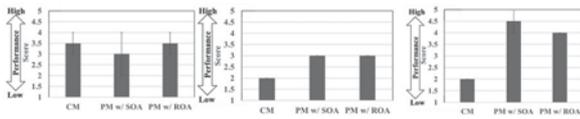


Fig.1: Experimental result for insect sound Fig.2: Experimental result for dental sound Fig.3: Experimental result for brake sound

### 1-4P-6

#### 1-4P-6 頭皮上に呈示した骨伝導音知覚に及ぼす耳栓装用の効果

Effects of earplugging on perception of bone-conducted sound presented onto the scalp

☆信夫大志, 大塚 翔, 中川誠司(千葉大)

- ◆機械工場等の強大騒音下では作業者は耳栓を装用することが多いが、円滑なコミュニケーションが損なわれてしまう。
- ◆一方、骨伝導音は耳栓装用下でも明瞭に聴取される。労働者が装着するヘルメットに骨伝導スピーカを内蔵させて頭皮上から呈示することで、強大騒音下でも利用可能な音声コミュニケーション・デバイスが実現できる。
- ◆耳栓装用下/非装用下において、頭皮上を含む様々な部位から呈示された骨伝導音の単音節明瞭度試験を実施した。
- ◆いずれの呈示部位においても、耳栓装用により明瞭度の上昇が観察された。耳栓効果(耳閉効果)の影響と考えられる。
- ◆耳栓装用の有無にかかわらず、従来部位の顎状突起において正答率が最も低下した。また、耳栓装用下では一部の頭皮上部位(Fpz, Cz)で乳様突起と同等以上の正答率が得られた。
- ◆呈示部位により、骨伝導音の主たる伝搬経路が変化し、明瞭度や耳栓効果の大小に影響を及ぼすことが確認された。

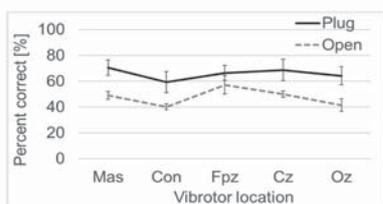


Fig. 1 Percent correct of the mono-syllable articulation test.

### 1-4P-5

#### 1-4P-5 自閉スペクトラム症者における両耳間時間差・レベル差の弁別閾値

Sensitivity to interaural time and level differences in individuals with autism spectrum disorder

○藤平晴奈(NTT CS 研), 糸井千尋(中央大学), 古川茂人(NTT CS 研), 加藤進昌(昭和大学発達障害医療研究所), 柏野牧夫(NTT CS 研)

- ◆自閉スペクトラム症 (autism spectrum disorder: ASD) では、音の定位能力の低下などの聴覚特性を伴うことが知られている。
- ◆音の水平面の定位は、両耳間時間差 (ITD) とレベル差 (ILD) を手がかりとしている。
- ◆本研究では、ASD 群と control 群の ITD と ILD に対する感度を比較した。
- ◆その結果、ASD 群では control 群に比べて ILD の弁別閾値が高いこと、例外的に高い ITD 弁別閾値を有するサブグループが存在することが分かった。
- ◆これらは ASD 者における音の水平方向の定位能力の低下に、ITD と ILD に対する処理能力の低下が関与していることを示唆している。

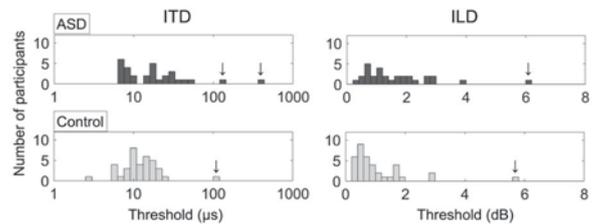


Fig. 1: Histograms of the thresholds for the ITD and ILD in the ASD and control groups. Arrows indicate the outliers.

### 1-4P-7

#### 1-4P-7 振動覚が骨伝導知覚に及ぼす影響の検討

Effects of vibration sensation on bone-conduction perception

☆富田繭子, 大塚翔, 中川誠司(千葉大)

- ◆概して 500 Hz 以下の低周波域の骨伝導音に対しては、聴覚に加えて振動覚が生じることがあり、正確な骨伝導聴力検査の妨げになっている。
- ◆本研究では、骨伝導によって生じる振動覚および聴覚の相互作用を明らかにすることを目的として、マスキング音呈示下の乳様突起における骨伝導検出閾を計測し、得られた結果から乳様突起における振動覚閾特性の推定を試みた。また、小指末節骨における振動覚閾特性と比較した (Fig.1)。
- ◆得られた結果から、骨伝導音を 60 dB HL 程度の強度で乳様突起に呈示すると、振動覚が生じる可能性が示唆された。また、小指末節骨が 200 Hz 付近を極小値とする U 字型の特性を示したのに対し、乳様突起における振動覚の周波数特性はほぼ平坦であった (Fig.2)。また、すべての周波数において乳様突起の振動覚閾は小指末節骨を上回った。これらの結果は、頭部では指先に比べて振動覚受容の必要性が小さいことに起因すると思われる。

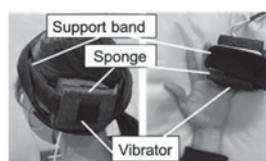


Fig. 1 Settings of the vibrator in the experiments.

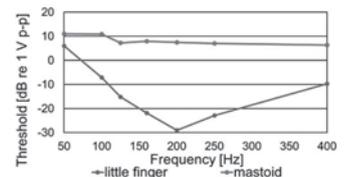


Fig. 2 Threshold of vibration sensation at the mastoid process of the temporal bone and the little finger.

### 1-4P-8

#### 1-4P-8 軟骨伝導における押付呈示圧および耳介形状変化が聞こえに及ぼす影響

Effects of static pressure and dimensions of the pinna on cartilage-conduction hearing

☆実川徹, 大塚翔, 中川誠司(千葉大)

- ◆骨伝導における振動子の装用性の改善を目的として“軟骨伝導”が提案され、外耳道閉鎖症用の補聴器やスマートホンに活用されている。
- ◆軟骨伝導では、振動子の接触面積や接触圧などに応じて聞こえが変化しますが、呈示条件による聞こえの差異が詳細に検討された例はない。
- ◆本研究では、軟骨伝導における耳介への押付呈示圧を変化させながら、耳介軟骨に生じる振動を計測した。また、耳介の形状・サイズと耳介に生じる振動の関係から、軟骨伝導の末梢伝搬過程の変化を推定した。
- ◆被験者1および2では、呈示圧が2Nに達すると、振動の周波数特性に変化が生じた(Fig.1)。また、耳輪と頭部との距離が比較的大きい被験者2に比べ、“耳輪と頭部との距離”および“耳介接続部と耳介長の比”の値が小さい被験者1では振動特性の変化が激しい。
- ◆今回の結果からは、耳介の形状パラメータのうち、“耳介接続部と耳介長の比”が呈示圧の上昇による振動特性の変化に特に作用していることを示唆している。

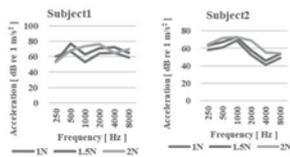


Fig.1 Acceleration of pinna for each subject.

### 1-4P-10

#### 1-4P-10 可聴周波数帯における遠位呈示骨伝導音の知覚特性 -時間分解能の検討-

Temporal resolution of distantly-presented bone-conduction hearing in the audible-frequency range

☆石川大夢, 大塚翔, 中川誠司(千葉大)

- ◆従来の骨伝導デバイスの装用性の問題を解決するために、骨伝導音を頭部から離れた位置(遠位)に呈示する方法(遠位呈示方式)が提案されている。しかしながら、可聴周波数帯の骨伝導音を遠位呈示した場合は骨伝導成分のみの評価が困難であり、知覚特性を検討した例は限定的である。
- ◆前報では、気導成分を十分に遮蔽した状態で、可聴周波数帯における遠位呈示骨伝導音の検出閾及び周波数弁別閾(DLF)計測を行い、伝搬距離の増大に伴い骨伝導音の減衰量が上昇すること、体内伝搬に伴う周波数情報の劣化が十分に小さいことを示した。
- ◆本報では、可聴周波数帯における遠位呈示骨伝導音の時間分解能を評価するために、前報と同様に気導成分を十分に遮蔽した状態で、時間変調伝達関数(TMTF)計測を行った。
- ◆搬送波周波数に関わらず、全ての変調周波数において呈示部位間の有意な差は確認されなかったことから、可聴周波数帯における遠位呈示骨伝導音の体内伝搬に伴う時間情報の劣化が十分に小さいことが示唆された。

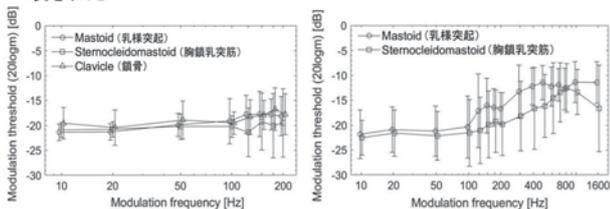


Fig.1: Modulation-detection threshold at carrier frequencies of 1000 Hz (left) and 4000 Hz (right)

### 1-4P-9

#### 1-4P-9 遠位呈示骨導音の体内伝搬過程における周波数特性の検討

Elucidation of frequency characteristics of propagation of distantly-presented bone-conducted sounds in the human body

☆土井公一朗, 大塚翔, 中川誠司(千葉大)

- ◆骨導超音波(骨導呈示された20kHz以上の高周波音)は明瞭な音として知覚可能であり、振幅変調(AM)することで音声伝達も可能である。
- ◆骨導超音波は首や体幹、上肢等の頭部から離れた部位(遠位)に呈示した場合でも知覚される。骨導超音波を遠位呈示した場合、振動の伝搬経路の形状や媒質が長く複雑になり、体内伝搬過程の周波数特性が変化するため、音質や音声の明瞭性が変化する可能性がある。
- ◆体内伝搬過程の周波数特性を明らかにするため、可聴周波数帯の骨伝導音を遠位呈示した際の聴覚受容器の近傍(外耳道)の振動を測定した。
- ◆胸鎖乳突筋や鎖骨は、低周波域(約200Hz以下)および高周波域(約2kHz以上)で高い振動の伝搬効率を示した(Fig.1)。
- ◆振動の伝搬経路となる部位の弾性や構造の違いが、振動の周波数特性を変化させることが示唆された。

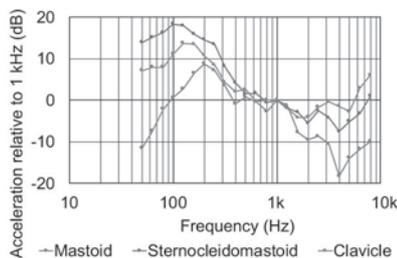


Fig.1: Frequency characteristics of the acceleration caused by distantly-presented bone-conducted sounds measured at the external auditory meatus.

### 1-4P-11

#### 1-4P-11 軟骨伝導における振動子呈示部位と末梢伝搬過程の関係

Relationships between transducer placements and peripheral propagation process of the cartilage conduction

☆菅原優聖, 大塚翔, 中川誠司(千葉大)

- ◆骨伝導の装用にかかる問題点を解決する方法として“軟骨伝導”が提案され、外耳道閉鎖症用の補聴器やスマートホンに活用されている。
- ◆軟骨伝導では気導成分と外耳道内放射成分が支配的とされているが、軟骨伝導デバイスの使用時には軟骨のみならず頭骨部にも振動面の一部が接触する場合がある。そのような接触部位の変化と伝搬成分や知覚の関係は明らかにできていない。
- ◆本研究では呈示部位ごとの骨導成分の評価を目的として、耳栓で気導成分を、外耳道内への注水によって外耳道内放射成分を遮断した際の聴覚閾計測と、レーザドップラ振動計による頭部表面振動の計測を行った。
- ◆周波数に関わらず、乳様突起よりも耳介における聴覚閾の上昇量が大きくなった(Fig.1)。外耳道内への注水により、外耳道伝搬成分が遮断されたためと考えられる。また、この結果は耳介呈示時に小さく、乳様突起呈示時に大きい頭部表面振動計とも矛盾しない(Fig.2)。
- ◆これらの結果は、振動面が耳介軟骨のみに接触する場合は、外耳道内に生じる成分が増大し、頭骨部に接触する条件では慣性骨導成分および圧縮骨導成分が発生することを改めて示している。

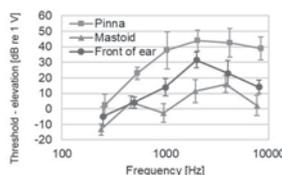


Fig. 1 Threshold-shifts at 40% water injection from 0% water injection

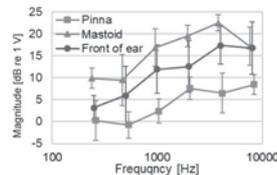


Fig. 2 The head vibration for each placement

## 1-4P-12

1-4P-12

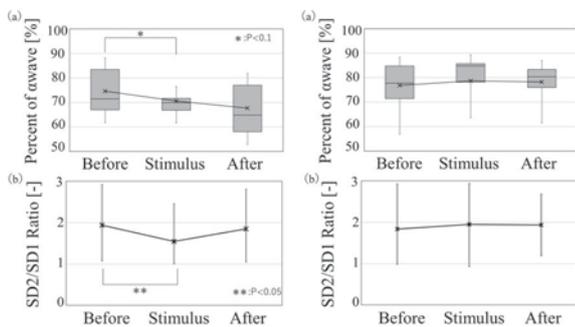
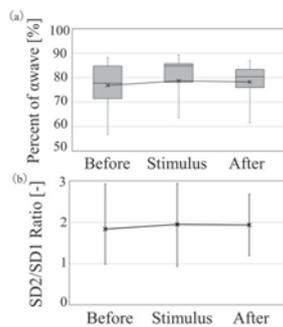
生体反応を用いた

## 聴覚刺激時のストレス評価\*

Stress evaluation by physiological measurement during auditory stimulation

☆中村祐太, 藤井崇史, 朝倉巧(東京理科大)

- ◆聴覚刺激に対するストレスを定量的に評価することを目的とし、脳波と心拍を用いて聴覚刺激が被験者に与えるストレスの定量化を行う。
- ◆本検討で使用した音源はホワイトノイズと川のせせらぎ音である。実験は聴覚以外の刺激を遮断するため簡易暗室で行い、閉眼・座位状態で無音5分、聴覚刺激5分、無音3分の順に計13分間で測定を行った。
- ◆音源を提示した際の脳波と心拍を測定し、脳波は $\alpha$ 波の比率をストレスの指標とした。心拍はポアンカレプロットによりSD2/SD1を算出しストレスの定量化を行った。

Fig.1: Presenting white noise  
(a)EEG, (b)ECGFig. 2: Presenting sound of river  
(a)EEG, (b)ECG