

### 2-1-1

#### 2-1-1 識別的変分自己符号化器学習による 特定話者モノラル音声分離

Single-channel Multi-speaker Separation via Discriminative Training of Variational Autoencoder Spectrogram Model

☆村島允也[1], 亀岡弘和[2], 李莉[1], 関翔梧[2], 牧野昭二[1]

[1]筑波大学

[2]日本電信電話株式会社 NTT コミュニケーション科学基礎研究所

- ◆本稿ではモノラル複数話者音声分離を扱う。近年深層クラスタリング(DC)をはじめとする識別的アプローチによる音源分離手法は非常に強力であることが報告されている。
- ◆一方で非負地行列因子分解(NMF)をはじめとする生成的アプローチによる音源分離は、学習条件とテスト条件のミスマッチに柔軟に対応できる点で識別的手法よりも優位になりうる。また NMF の手法には、ウィーナフィルタ出力と学習サンプルのスペクトログラムの誤差を規準として基底学習を行う識別的 NMF が提案されている。NMF に基づく手法は特定の音源に対して有効である一方、線形表現が難しい音源に対しては分離性能が低下する問題があった。
- ◆近年ではDNNによる非線形表現能力を利用した変分自己符号化器(VAE)による生成的アプローチの手法が多く提案されている。亀岡らはCVAE音源モデルによる優決定条件下での音源分離手法MVAE法を提案している
- ◆本稿ではCVAE音源モデルを、モノラルの音源分離に応用した変分自己符号化器による音源分離(VASS)を提案し、その最適化アルゴリズムをEMアルゴリズムに基づき示す。さらにVAEモデルとNMFモデルの関係性に着目し、識別的NMFで展開される識別的学習をVASSに取り入れた識別的VASSを提案する。本稿では二話者混合信号に対して、提案手法2つと各種従来法との分離精度の比較を行いその有効性を示す。

### 2-1-3

#### 2-1-3 深層学習に基づく周波数帯域補間手法 による音源分離処理の高速化

Fast audio source separation based on deep-neural-network-based frequency component interpolation

☆渡辺瑠伊, 北村大地(香川高専), 猿渡洋(東大), 高橋祐, 近藤多伸(ヤマハ)

目的:

多チャンネル音源分離(MASS)において、深層学習(DNN)を用いて高速化を行う音源分離フレームワークを提案

方法:

一部の周波数ピンを間引き、残った周波数ピンにのみ MASS を適用  
分離しなかった周波数ピンは DNN により分離信号を別途予測  
本稿では Fig. 1 に示す 2 種類の周波数ピンの間引き方法を提案

結果:

提案フレームワークにおいて 3 種類の間引き方法で実験  
一部の条件において分離精度の改善を確認

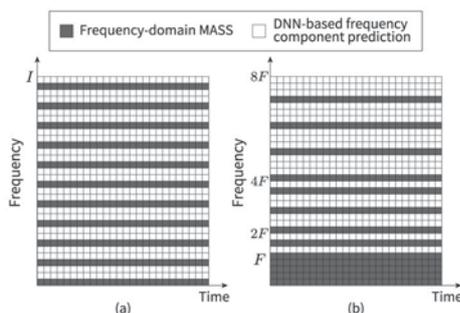


Fig. 1: Proposed frequency decimations: (a) even-interval and (b) uneven-interval decimations.

### 2-1-2

#### 2-1-2 マルチモーダル音源分離のための モーダル間対応学習

Cross-modal learning for audio-visual audio source separation

◎牧島直輝, 庵愛, 田中智大, 高島瑛彦, 折橋翔太, 増村亮(NTT)

- ◆話者の顔の動画情報などを補助入力として混合信号から混合前の音声信号を推定するマルチモーダル音源分離では、従来、分離信号と教師信号の誤差のみを用いてモデルを学習していた。
- ◆従来の学習では、分離信号に含まれる音声の特徴や話者性といった要素を考慮していないため、妨害音の残留や歪みが生じることがある。
- ◆本稿では、これを改善するため、映像信号と分離信号を対応させるモーダル間対応学習を提案する。
- ◆評価実験により、提案法による学習で分離性能が向上することを示す。

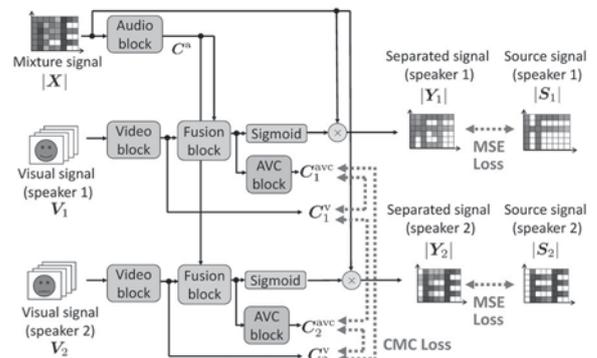


Fig. 1: Overview of proposed method.

### 2-1-4

#### 2-1-4 経験ベイズ独立深層学習行列分析による 多チャンネル音源分離

Multichannel audio source separation by empirical Bayesian independent deeply learned matrix analysis

☆蓮実拓也, 中村友彦, 高宗典玄, 猿渡洋(東大), 北村大地(香川高専), 高橋祐, 近藤多伸(ヤマハ)

- ◆本稿では、独立深層学習行列分析を拡張し、経験ベイズ法に基づき学習された DNN を用いる、経験ベイズ独立深層学習行列分析(empirical Bayesian IDLMA: EB-IDLMA) を提案する。
- ◆提案法では、DNN によって得られた分離信号の分散推定値の信頼度を考慮した分離フィルタ推定を行う。
- ◆これを実現するため、分離信号の分散を潜在変数として扱い、当該潜在変数に対する事前分布のハイパーパラメータについて DNN で推定する。
- ◆楽音分離実験により分離性能を評価し、信頼度を推定することの有効性を確認した。

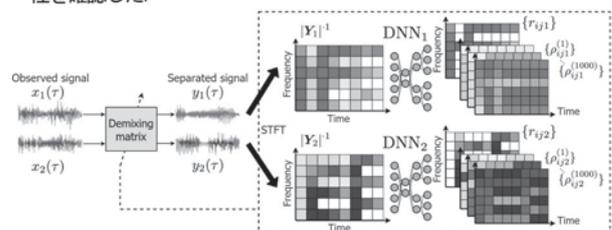


Fig. 1: Separation process of proposed EB-IDLMA.  $r_{ijn}$  is variance and  $\rho_{ijn}^{(k)}$  is weight of anchors for degree of freedom in Student's t distribution.

2-1-5

2-1-5 SepNet: 高速多チャンネル音源分離のための分離行列予測ネットワーク

SepNet: DNN-based demixing matrix predictor for determined multichannel source separation

☆井上翔太(筑波大), 亀岡弘和(NTT CS 研), 李莉(筑波大), 牧野昭二(筑波大)

- ◆決定条件下での音源分離手法として対数尤度の最大化を基準とした反復アルゴリズムが研究されている。分離行列の更新規則として用いられる IP 法は少ない反復での収束を可能とするが、反復当たりの計算コストは音源数の4乗に比例して増加する。
- ◆本手法では、DNN ベースの多チャンネル観測信号からの分離行列手法である SepNet を提案する。
  - パーミュテーションフリーな分離行列推定の過程の獲得を目的として、従来の反復最適化の枠組みを参考にした DNN を設計し、さらにモデル学習の損失関数に PIT を導入した。
  - ネットワークの各層に CNN を用い、フォワード計算の計算コストを音源数の2乗に比例する増加に留めた。
- ◆2音源および3音源の音声分離実験を実施した。提案法は従来法より高速かつ高精度な分離行列の推定が可能であることが示された。

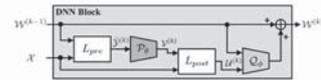
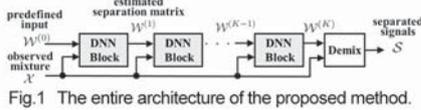


Fig.2 A block diagram for each DNN block.

2-1-7

2-1-7 要素選択による次元削減を用いた DNN 音声強調処理の低演算量化

A low-computational DNN-based speech enhancement based on sample point selection

☆春田智穂, 小野順貴(都立大)

- ◆難聴者の雑音下での言葉の聞き取りに対して、**DNN を用いた音声強調処理**が高い効果をあげており、**補聴器**への応用が期待されている。
- ◆一方で補聴器には演算量の制約があり、乗算の回数が多いと大きな計算負担となる。よって大規模な DNN は搭載することができない。
- ◆本研究では同処理の**低演算量化**の新たなアプローチとして、**要素選択による次元削減**を行ってから DNN で音声強調を行う手法を検討する。
- ◆本手法において次元削減は元のデータからの要素の選択のみで行われるため、乗算は発生しない。
- ◆本稿では、提案手法により次元削減を行った場合と行わなかった場合との音声強調の性能を比較した。

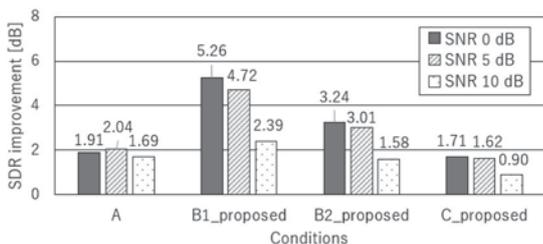


Fig. 1: SDR improvement of processed noisy speech (café noise). A: conventional ( $N^2$  multiplications), B1: proposed ( $N^2$  multiplications), B2: proposed ( $N^2/2$  multiplications), C: proposed ( $3N^2/8$  multiplications).

2-1-6

2-1-6 Teacher-Student学習を用いた Wave-U-Netによる低遅延リアルタイム音声強調

Wave-U-net with teacher-student learning for low latency online speech enhancement

☆中岡想太郎, 李莉, 井上翔太, 牧野昭二(筑波大)

- ◆Wave-U-NetはEnd-to-End型のモノラル音源分離手法であり、時間領域で動作するため高品質な位相情報をモデル化し、音声分離や音声強調のタスクにおいて高い性能を示している。
- ◆従来のWave-U-netはオフライン処理の手法だが、時間領域で動作するためSTFTにおける窓長の制約を持たず、低遅延のオンライン処理に拡張できると考えられる。
- ◆本稿では、低遅延を満たすための入力セグメント長の短縮による性能低下を防ぐために、Teacher-Student学習を適用したオンラインWave-U-Netを提案する。
- ◆実験により、提案手法はシステム遅延を約6.7msに留めながら音声強調性能を向上させ、低遅延かつ高い音声強調性能を達成可能である事を示した。

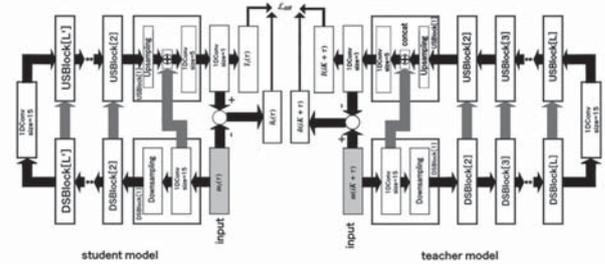


図1: Teacher-Student 学習に用いられるオフラインおよびオンライン Wave-U-net モデルのネットワーク構造と学習法。

2-1-8

2-1-8 音場補間を用いた円状マイクロホンアレイの回転に頑健なビームフォーミング

Beamforming robust for rotation of circular microphone array with sound field interpolation

◎若林佑幸, 小野順貴(東京都立大)

- ◆多くのアレイ信号処理は音源とマイクの位置(系)の時不変性を仮定しているが、実応用ではこの仮定は常に満たされず、系が変化する度にフィルタ更新などの複雑な演算を繰り返し行う必要がある。
- ◆我々はこの問題に対し、前回の研究発表会において、円状マイクロホンアレイ(CMA)の対称性や円周上の音場の周期性に着目した回転移動に対する音場補間手法を提案した。
- ◆本研究では、ビームフォーマ(BF)を事前に設計した後に系が変化したとき、提案手法により音場を補間し、事前設計BFを更新せずに適用することで系の時不変性を仮想的に実現できることを確認する。
- ◆最小パワー無歪応答(MPDR)BFを用いたシミュレーションにより、CMAが回転移動したとき、提案手法(Int)によりSDRやSIRが改善し、時不変性を仮想的に実現できることを確認した(Fig.1参照)。

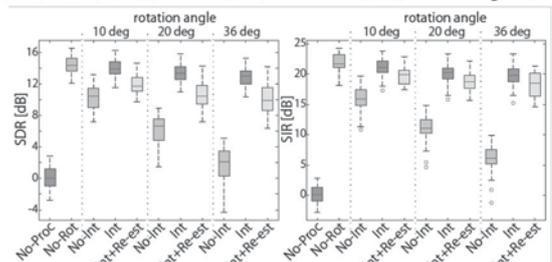


Fig. 1: SDR and SIR for five cases: no rotation (No-Rot), w/o interpolation when CMA rotates (No-Int), w/ interpolation (Int), re-estimation of BF after interpolation (Int+Re-est).

### 2-1-9

#### 最適化によるパルス符号変調信号の再量子化処理の高速化

Fast algorithm of optimization-based re-quantization for PCM signal

☆金原翔, 米谷昭彦(工大・工)

- ◆量子化ノイズの発生対策として、従来扱われている $\Delta\Sigma$ 変調器では、一般的にノイズの総量が増えてしまうため、著者らは複数ステップにわたる量子化ノイズに周波数重みを付けて、最小にする方法を提案している。しかし、最適化計算の負荷が大きくなってしまったといった問題点があるため、目標を組み込み用途のリアルタイムでの計算処理と定め、その高速計算手法を提案するとともに DSP 実装を行い、実機評価した。
- ◆高速計算手法として、Look-up table(LUT)を用いたアルゴリズムを開発した。従来の最適化における直接探索を省略することができたが、LUT のためのメモリ容量を必要としてしまう。LUT サイズを抑えるためのノウハウについても言及した。最適評価のステップ幅が4であるとき、LUT は8192語(2bit/語)のサイズで所望の特性が得られることを明らかにした。
- ◆テレビの音声データを想定し、実機検証を行った。最適評価のステップ幅を4とし、Fs = 48 kHz で16ビットに再量子化するプログラムを DSP の開発評価ボードに実装した。5次のノイズシェーピング特性を持たせたとき、1 sample のデータを処理するのに要する時間は約5 $\mu$ sであり、組み込み用途においても実用的であることが実証された。

### 2-1-11

#### 2-1-11 非凸最適化によるスパース無歪ビームフォーマ

Sparse distortionless beamformer based on non-convex optimization

☆河村泰雅(徳山高専), 矢田部浩平(早稲田大), 宮崎亮一(徳山高専)

- ◆MVDR を改良し、残留成分を減少させるスパースな出力を得ることが可能な、 $\ell_1$ ノルム最小化によるビームフォーマが提案されている。
- ◆ $\ell_1$ ノルムを使用することで信号全体に不利なバイアスが掛かってしまう問題点がある。バイアスが掛からない $\ell_0$ ノルムを使用することが理想的だが、最適化が困難である。
- ◆バイアスが小さく最適化し易い非凸最適化によるビームフォーマを提案する。従来手法より SDR が高いことを確認した (Fig. 1)。

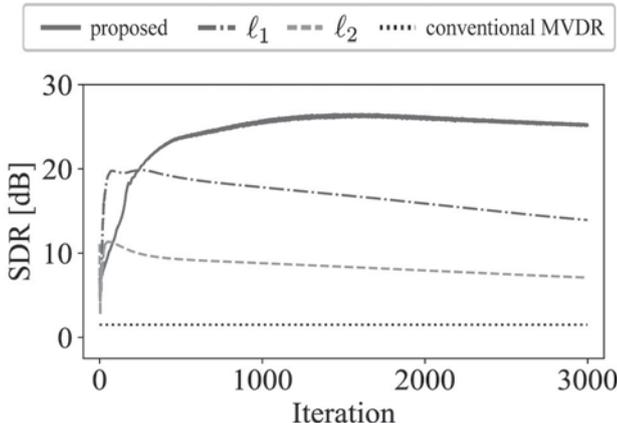


Fig. 1: SDR for each iteration.

### 2-1-10

#### 2-1-10 音源分離のための一般化 KL ダイバージェンスに基づいた位相復元

Phase reconstruction using the generalized KL divergence for monaural source separation

◎升山義紀, 矢田部浩平, 長友健人, 及川靖広(早大理工)

- ◆ STFT 係数の振幅のみを操作する音源分離のための位相復元
  - 振幅から適切な位相(または時間領域信号)を推定し、分離音の品質を改善
- ◆ 従来手法
  - 時間領域信号の推定で、振幅の近さとしてユークリッド距離の2乗を利用
    - ⇒ 元音源の振幅が小さい時間周波数ビンで誤差が大きくなり耳障り
  - 最適化アルゴリズム内で振幅の分離結果を常に維持
    - ⇒ 振幅の分離結果に含まれる誤差が復元結果に与える影響大
- ◆ 提案手法
  - 振幅の近さとして、一般化 KL ダイバージェンスを利用
    - ⇒ 元音源の振幅が小さい時間周波数ビンでの誤差を削減
  - 最適化アルゴリズムとして交互方向乗数(ADMM)を利用
    - ⇒ 時間領域信号と同時に分離結果の振幅も適応的に変化
- ◆ 雑音除去での数値実験
  - 振幅の推定精度が良い条件 (IRM, tIAM) で、提案手法の有効性を確認

Table 1: Average SI-SDR of reconstructed 200 speech signals.

	SI-SDR					
	SNR = 0 dB			SNR = 10 dB		
T-F Mask	SS	IRM	tIAM	SS	IRM	tIAM
Observed phase	4.88	11.89	11.90	14.36	18.35	18.99
GLA	4.68	12.32	14.06	14.05	18.10	21.16
MISI	4.82	13.12	17.89	14.30	19.24	24.15
CWF	<b>6.30</b>	13.01	13.45	<b>14.97</b>	19.46	19.68
Proposed	5.17	<b>13.63</b>	<b>18.72</b>	14.41	<b>19.97</b>	<b>25.56</b>

### 2-1-12

#### 2-1-12 Noisy-target Training: DNN 音声強調におけるクリーン音声を必要としない学習法

Training of DNN for speech enhancement without clean speech

◎藤村拓弥(徳山高専), 小泉悠馬(NTT), 矢田部浩平(早稲田大), 宮崎亮一(徳山高専)

- ◆雑音が混入した音声のみから音声強調用の DNN を学習する新しい枠組み NyTT (Noisy-target Training) を提案する。
- ◆NyTT では雑音混入音声に更に雑音を加えた信号から元の雑音混入音声を予測するように学習する。
- ◆ 実験では、(i) 提案法がクリーン音声なしで学習でき (Table 1), (ii) 学習データと評価データにミスマッチがある場合では、クリーン音声をを用いる手法 (CTT, NeTT) と同程度の性能が得られ (Table 2), (iii) SNR が15 dB 以上の音声はクリーン音声として学習に利用できることが明らかになった (Fig. 1)。

Table 1: Results on VoiceBank-DEMAND (no mismatch of training and testing datasets). Table 2: Results on TIMIT-MOBILE (with mismatch of training and testing datasets).

Method	SI-SDR	PESQ	CSIG	CBAK	COVL	Method	SI-SDR	PESQ	CSIG	CBAK	COVL
Input	9.21	1.97	3.35	2.44	2.63	Input	4.69	1.30	2.73	1.75	1.94
CTT	<b>19.53</b>	<b>2.68</b>	<b>3.83</b>	<b>3.37</b>	<b>3.25</b>	CTT	<b>12.60</b>	<b>2.02</b>	<b>3.22</b>	<b>2.71</b>	<b>2.58</b>
NeTT	19.50	2.63	3.77	3.34	3.19	NeTT	12.26	1.99	3.13	2.67	2.52
NyTT	17.66	2.30	3.19	3.01	2.72	NyTT	12.09	1.95	3.41	2.61	<b>2.64</b>
NyTT (L)	17.72	2.31	3.23	3.02	2.75	NyTT (L)	12.38	1.91	<b>3.43</b>	2.58	2.63

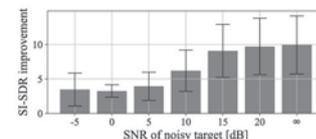


Fig. 1: Relationship between SNR of noisy signal utilized in NyTT and its SI-SDR improvement.

## 2-1-13

### 2-1-13 補助関数法に基づく幾何学的制約付き独立ベクトル分析の車室内音声強調への適用

Study on geometrically constrained IVA with auxiliary function approach and VCD for in-car communication

☆後藤加奈, 高橋理希, 李莉, 牧野昭二, 山田武志(筑波大学)

- ◆車室内コミュニケーションの円滑化のための音声強調システムを検討する。車室内でのマイクと搭乗者の相対位置は一定であるため、指向性を形成する音源分離手法が有用である。
- ◆GCAV-IVAは幾何学的制約付き独立ベクトル分析(GCIVA)に補助関数法を組み合わせた手法で、音源やマイク位置の空間情報を幾何学的制約として事前に与え、独立ベクトル分析(IVA)の最適化問題と組み合わせた手法である。幾何学的制約によって分離後の目的音声を選択する後段処理が不要となり、補助関数によって計算の収束が早く、ステップサイズの調整が不要になるという利点を持つ。先行研究ではシミュレーション状況下での有効性が確認されている。
- ◆本研究ではGCAV-IVAによる車室内音声強調システムへの適用を検討するためマイク・音源数を3とし評価実験を行った。テストデータは車室内で測定されたインパルス応答を用いて作成した。またGCAV-IVAを、DOAを与えた固定制約と、訓練データからステアリングベクトルを推定するデータ適応制約に適用した。実験の結果、GCAV-IVAは従来手法より高い性能を示した。

Methods		SDR	SIR	SAR
BSS	AuxIVA	16.32	22.76	17.66
Fixed	DS	0.14	0.15	31.13
	GCAV-IVA(1)	8.06	12.78	11.30
	MVDR	17.22	29.13	17.61
Data-adaptive	maxSNR	19.09	27.33	20.13
	GCAV-IVA(2)	19.12	26.16	20.25

Average SDR, SIR, and SAR [dB] over test dataset achieved by different method

## 2-1-15

### 2-1-15 スタガードモデル化三重対角型周波数共分散行列を用いた独立半正定値テンソル分析によるブラインド音源分離

Independent positive semidefinite tensor analysis using staggered-model-based tridiagonal frequency-covariance matrix for blind source separation

◎近藤樹(東大), 高宗典玄(東大), 北村大地(香川高専), 猿渡洋(東大), 池下林太郎(NTT), 中谷智広(NTT)

- ◆独立半正定値テンソル分析(IPSDTA) [Ikeshita, 2018] [Kondo+, 2020]では、計算コストの観点から周波数共分散行列に対してブロック対角制約を課していたが、異なるブロックの周波数との相関を考慮できないという問題点があった。
- ◆周波数共分散行列を三重対角行列としてそれをCholesky分解によりモデル化するIPSDTA [近藤+, 2020]が提案されたが、その性能は限定的であった。
- ◆本稿では三重対角行列を2つの互い違いのブロック対角行列の和として表現するスタガードモデル化を提案する。
- ◆音源分離実験により性能を評価し、その性質について考察する。

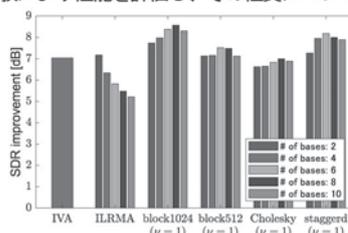


Fig.1: SDR improvement for each method.

## 2-1-14

### 2-1-14 車室内の三角マイクロフォンアレイへのヴァーチャルマイクロフォン技術の適用

Applying Virtual Microphones to Triangular Microphone Array in in-Car Communication

☆瀬川華子, 高橋理希, 李莉, 陣在遼河, 牧野昭二, 山田武志(筑波大学)

- ◆近年、情報通信技術を用いた車室内におけるコミュニケーション補助システムの開発が検討されている。
- ◆車室内においては音声の到来方向が予測しやすく、ビームフォーミングにおける音声強調が効果的であるが、劣決定条件になりやすい車室内では十分な強調性能が得られない。
- ◆実マイクの信号から仮想的な信号を合成するヴァーチャルマイクロフォン技術を用いて劣決定条件におけるビームフォーミングの強調性能が改善することが示されている。
- ◆本研究では、ヴァーチャルマイクロフォン技術の外挿とWPVM配置を劣決定条件における車室内に設置された三角マイクロフォンでの音声強調に適用し、実際に測定したインパルス応答を用いた実験を行い、劣決定条件下での車室内の三角マイクロフォンアレイにおいても音声強調性能を改善するためにヴァーチャルマイクロフォンが有効であることを示した。

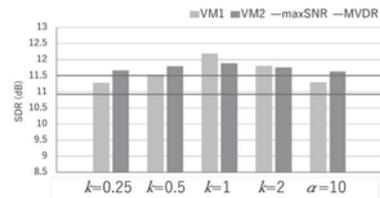


Fig.1 result of driver's speech enhancement(window close)

## 2-1-16

### 2-1-16 多変量複素 Sub-Gauss 分布に基づく同時対角化制約付き多チャンネル非負値行列因子分解における majorization-equalization アルゴリズムを用いた更新則

Majorization-equalization-algorithm-based update rules of joint-diagonalizability-constrained multichannel nonnegative matrix factorization based on multivariate complex sub-Gaussian distribution

◎加茂茂佳, 久保優騎, 高宗典玄(東大), 北村大地(香川高専), 猿渡洋(東大), 高橋祐, 近藤多伸(ヤマハ)

- ◆空間相関行列の同時対角化制約を設け、多チャンネル非負値行列因子分解(MNMF)の生成モデルをsub-Gauss分布へ一般化した同時対角化制約付きsub-Gauss MNMFの更新則は、majorization-minimization(MM)アルゴリズムに基づき導出された [Kamo+, 2020].
- ◆本稿では、同時対角化制約付きsub-Gauss MNMFの新たな更新則をmajorization-equalization (ME) アルゴリズムを用いて導出する。
- ◆パラメタ間の収束速度の違いが音源分離性能に影響すると報告されているため [Mitsui+, 2017], パラメタごとにMMまたはMEアルゴリズムを選択し組み合わせを変化させることで性能改善を期待できる。
- ◆音源分離実験により、新たな更新則導入の有効性を示す。

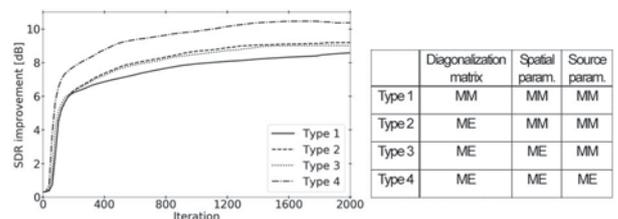


Fig.1: SDR behaviors of four types of sub-Gauss MNMF.

	Diagonalization matrix	Spatial param.	Source param.
Type 1	MM	MM	MM
Type 2	ME	MM	MM
Type 3	ME	ME	MM
Type 4	ME	ME	ME

## 2-1-17

### 2-1-17 低遅延でオンライン動作する残響除去と音源分離の同時最適化

Joint optimization of source separation and dereverberation that operates with online and low latency

☆上田哲也(筑波大), 中谷智広, 木下慶介, 池下林太郎, 荒木章子 (NTT), 牧野昭二 (筑波大)

- ◆音源分離手法の1つである独立ベクトル分析(IVA)を補聴器や車室内に適用するには低遅延化とオンライン化が求められる。
- ◆オンライン動作するIVAを低遅延化するために短時間フーリエ変換の窓長を短くすると、残響によって分離性能が劣化する。これに対し、オンラインIVAの前処理としてオンライン動作する残響除去(WPE)手法を結合した手法(WPE+IVA)は窓長を削減しつつ分離性能の改善が期待できるが、IVAとWPEの同時最適化を保証していない。
- ◆本稿では、IVAとWPEの同時最適化を保証したバッチ処理での手法をオンライン化した手法を提案する。車室内で収録したインパルス応答と走行雑音を用いた音源分離実験により、提案手法の有効性を評価する。

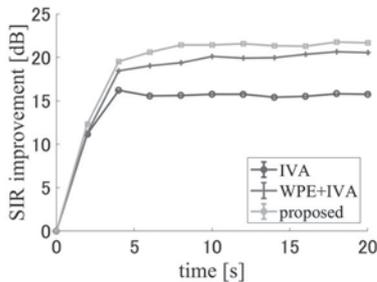


Fig.1: Separation performance in comparison with each online method.

## 2-2-2

### 2-2-2 Real-time Neural Machine Speech Chain

○ Sashi Novitasari<sup>1</sup>, Andros Tjandra<sup>1</sup>, Tomoya Yanagita<sup>1</sup>, Sakriani Sakti<sup>1,2</sup>, Satoshi Nakamura<sup>1,2</sup> (<sup>1</sup>NAIST, <sup>2</sup>RIKEN AIP)

Machine speech chain framework enables the semi-supervised construction of the end-to-end ASR and TTS systems by connecting them through a feedback loop that makes them improve together. This framework was inspired by the mechanism of listening while speaking in humans for speech improvement. However, the systems require a long delay for output and feedback due to the global attention mechanism. In this work, we constructed an incremental machine speech chain framework to reduce the delays in the original machine speech chain. The incremental speech chain connects incremental ASR (ISR) and incremental TTS (ITS) using a short-term feedback loop. Our experiment results show that the proposed framework reduced the delay required in the original framework while keeping a close performance to it.

## 2-2-1

### 2-2-1 Improving ASR with Multimodal Machine Chain

Johanes Effendi<sup>1,2</sup>, Andros Tjandra<sup>1,†</sup>, Sakriani Sakti<sup>1,2</sup>, Satoshi Nakamura<sup>1,2</sup> (<sup>1</sup>NAIST, <sup>2</sup>RIKEN AIP)

Gathering a substantial amount of parallel data is one of the major problems in building a multimodal model. Previously, a machine speech chain was proposed to enable the training of ASR and TTS to assist each other in semi-supervised learning. It successfully avoids the need for a large amount of paired speech and text data by using a speech or text only unpaired data in a closed-loop mechanism. However, this work has not yet addressed visual modality, although visual modality is also one of the important senses in human communication.

In this study, we formulate a multimodal model collaboration between automatic speech recognition (ASR), text-to-speech synthesis (TTS), image captioning (IC), and image generation (IG), in which they support each other under a weakly-supervised chain training strategy. The results reveal that with the help of IC and IG, ASR and TTS can still be improved using an image-only dataset.

<sup>†</sup>This work was done when he was at NAIST, he is currently at Facebook AI, USA.

## 2-2-3

### 2-2-3 マルチタスク学習による方言識別を考慮した End-to-End 日本語方言音声認識

Multitask Learning for End-to-End Japanese Dialect Speech Recognition and Dialect Identification

☆今泉遼 (都立大), 増村亮 (NTT), 塩田さやか, △貴家仁志 (都立大)

- ◆データ量の少ない方言を高い精度で認識する手法の1つにデータ量の豊富な標準語データと方言データを混合することが挙げられる。
- ◆しかし日本語方言は標準語と異なるアクセントや文節の切れ目を持つ。そのため方言固有の情報が標準語の学習に不要なバイアスをかけてしまい、標準語の認識性能が下がる可能性がある。
- ◆標準語、日本語方言それぞれを高い性能で認識するために日本語方言識別と方言音声認識を同時に行うマルチタスク学習を提案する。ネットワーク構造を Fig.1 に示す。
- ◆従来の方言データと標準語データを単純に混合したモデル化手法と比較して提案法は CER が 1.1% 程度低下し、性能向上が見られた

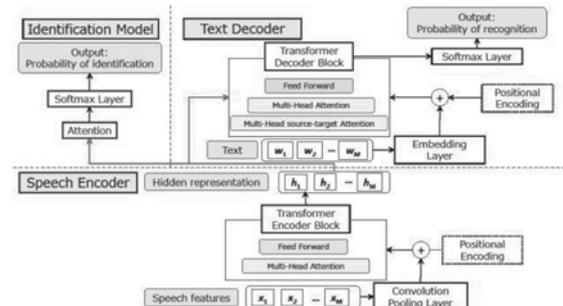


Fig.1 Network structure for multitask learning with end-to-end ASR and dialect identification

### 2-2-4

#### 2-2-4 聴覚障害者の音声データの収集と音素認識による評価 Collection of Deaf Speech and Evaluation by Phoneme Recognition

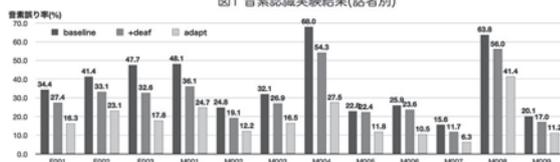
○小林 彰夫, 安 啓一(筑波技大), 西崎 博光(山梨大・院), 北岡 教英(豊橋技科大)

- ◇ 聴覚障害者の音声認識
  - 聴覚障害者のうち、手話を日常的に用いているのは概ね 25%
  - 音声による意思疎通を行う者が 50%以上 (補聴器・人工内耳装用と筆談の総計)
  - 障害者がコミュニケーションツールとして音声認識を使う可能性
    - 聴覚障害者は発話が不明瞭になりやすく、認識性能が劣化
    - 多数の聴覚障害者を対象とした認識性能の分析例はほとんどない
- ◇ 音素認識実験
  - 学習データ: 健聴者 (ASJ+JNAS), 聴覚障害者 12名 (バランス文)
  - 評価データ: 健聴者 3名, 聴覚障害者 12名 (バランス文 B/C セット)
  - ネットワーク: ESPNet (Hybrid Attention/CTC)
    - 健聴者の音声データのみで学習
    - 聴覚障害者音声により baseline をチューニング
    - B(C) セットで話者ごとにチューニング
  - 実験結果
    - 聴覚障害者の場合、CTC 単独で学習した方が音素誤り率が低い
    - ⇒ 不明瞭発話では CTC による系列への制約が有効
    - 音素誤り率は個人によりばらつきが大きい
    - ⇒ 話者によっては適応化しても誤りの削減が不十分

表1 音素認識実験結果

CTC weight	音素誤り率 (%)		
	0.0	0.5	1.0
hearing	10.6	9.0	10.4
deaf(baseline)	40.6	38.2	37.9
deaf(+deaf)	41.6	30.6	30.0
deaf(adapt)	31.8	18.6	18.3

図1 音素認識実験結果(話者別)



### 2-2-6

#### 2-2-6 ニューラル機械翻訳により推定された読み仮名・韻律記号を入力とする日本語 End-to-End 音声合成の評価

Evaluation of Japanese End-to-End Speech Synthesis Method Inputting Kana and Prosodic Symbols Predicted by Neural Machine Translator

☆懸川直人, 原直, 阿部匡伸(岡山大・HS 統合科学研), 井島勇祐(NTT)

- ◆ 本稿では、ニューラル機械翻訳により生成された読み仮名・韻律記号を入力とする日本語 End-to-End 音声合成システムの性能を評価した。
- ◆ 従来システムを用いて約 500 万文のニュース記事から読み仮名、韻律情報を推定し、これを正解として学習をおこなった。
- ◆ 読み仮名・韻律記号列の推定部分の有効性を確認するために、人手でラベリングされたテストセットでの評価を行い、読み仮名、アクセント句境界、アクセント核についてそれぞれ 98.0%, 92.8%, 89.8% の精度で正しく推定できることを確認した。
- ◆ 提案システムによる合成音声で主観評価実験を行い、アクセント誤りが 1 つ以上存在する場合、そのアクセント誤りの数によって MOS 値の有意な差が現れないことを確認した。



Fig. 1 MOS in the subjective evaluation of accent and intonation quality. The error bar indicates 95% confidence interval.

### 2-2-5

#### 2-2-5 Onoma-to-Wave: オノマトペからの環境音合成手法の提案

Onoma-to-Wave: Environmental sound synthesis using onomatopoeic words

☆岡本悠希(立命館大学), 井本桂右(同志社大学), 高道慎之介(東大院), 山西良典(関西大), 福森隆寛, 山下洋一(立命館大学)

- ◆ 本研究では **オノマトペを入力とする環境音合成手法を提案**
  - **オノマトペ**: 音の特徴を模倣した語 (例: 「ピー」など)
    - ↳ 音の特徴を表現するのに有効であるとされている
- ◆ 深層学習によりオノマトペと環境音の対応関係を学習
  - 系列変換モデル (sequence-to-sequence) により学習
- ◆ オノマトペに加え、音の種類の情報を入力とし、生成音を制御
  - 「笛の音」のような音の種類を表すラベルも入力することで、オノマトペだけでなく音の種類 (音響イベント) の制御を実現
- ◆ 主観評価実験により合成音の評価を実施
  - 生成音が入力オノマトペをどの程度表現できているかの評価において従来法 (KanaWave) より高いスコアを獲得

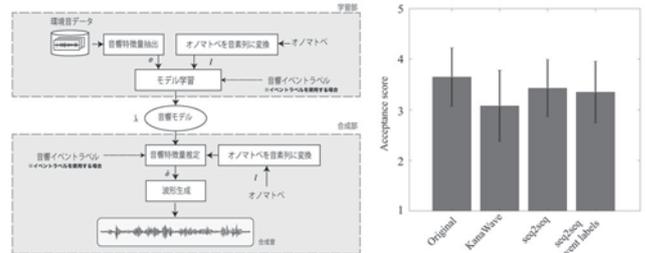


図1: 提案するオノマトペからの環境音合成手法の概要

図2: オノマトペに対する環境音の許容度

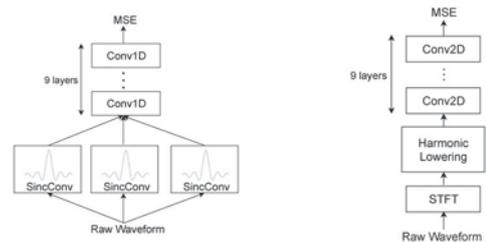
### 2-2-7

#### 2-2-7 音声の周波数特性を考慮した畳み込み層を持つ波形生成モデルの検討

A study of waveform generation model with convolutional layer considering frequency characteristics of speech

☆水田和輝(東大), 郡山知樹(東大院・情報理工), 猿渡洋(東大院・情報理工)

- ◆ 近年研究されている波形生成モデルは STFT ロスと呼ばれる周波数情報を用いた損失関数を加えることで効率良く音声の時間・周波数分布を学習することに成功している。したがって、モデルが能動的に周波数に注目して学習することが可能になれば、さらなる学習効率の向上が期待される。本論文は、周波数情報に注目する畳み込み層である sinc 関数をカーネルとする SincConv と音響信号の調和構造を捉える HarmonicConv を Fig.1 のように Discriminator に導入した音声波形生成モデルを提案する。これらの提案モデルによる合成音声の品質は早期の学習段階において既存のモデルによる合成音声の品質を上回ることを主観評価実験で示した。



(a) SincConv Discriminator

(b) HarmonicConv Discriminator

Fig.1: Discriminators of the proposed speech waveform generation models.

## 2-2-8

### 2-2-8 テンソル分解を用いた LPCNet に基づく 軽量なニューラルボコーダの評価

Evaluation of lightweight neural vocoder based on LPCNet with tensor decomposition

©金川裕紀, 井島勇祐(NTT)

- ◆WaveNet に代表されるニューラルボコーダは、従来の信号処理ベースのボコーダよりも高品質な音声を合成できるが計算量が大きく、高速な推論に GPU 等の並列プロセッサが必要であることが多い。
- ◆CPU での高速推論が可能なニューラルボコーダの一つに、計算量の小さい DNN とソースフィルターモデルを融合した LPCNet がある。しかし高速化の工夫は全体のうち、1 つの GRU のみに適用されており、残りの DNN モジュールが処理時間の約半分を占める。
- ◆本稿では高速化のため、処理時間が大きい残りの DNN モジュールをテンソル分解で低ランク近似。提案法によって、
  - CPU 上で 1.26 倍の高速化を実現
  - 上記速度の条件において、自然音声から抽出した音響特徴量と TTS で予測した音響特徴量の両方で、オリジナルの LPCNet から自然性を損なわず動作

Table 1: Average real time factors (RTFs) @ Intel Core i7-8750H 2.20GHz. "Speed enhancement" denotes the improvement over original LPCNet.

Method	RTF	Speed enhancement
Original LPCNet	0.35	-
HOSVD DualFC <sup>**1</sup>	0.33	1.08x
TTGRU <sub>b</sub> (v1) <sup>**2</sup>	0.30	1.16x
HOSVD DualFC +TTGRU <sub>b</sub> (v1)	0.28	1.26x

\*\*1: DualFC's parameter reduction by 82.5%.  
\*\*2: GRU<sub>b</sub>'s parameter reduction by 86.7%.

## 2-2-10

### 2-2-10 周期・非周期成分の分離に基づく ニューラルボコーダによる 音声波形のモデル化の検討

An investigation of modeling speech waveform by neural vocoder based on periodic/aperiodic decomposition

◎法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一 (名工大)

- 我々は、明示的な周期信号と非周期信号を入力とし、並列型や直列型のモデル構造を用いることで、周期・非周期成分が自動的に分離される形で音声波形をモデル化できる手法を提案した。
- 本稿では、並列型のモデル構造を持ったボコーダ (Fig. 1) において、事前に音声波形を周期・非周期成分に分離しモデル化した場合 (Fig. 2) と比較することで、学習の過程で適切に周期・非周期成分に分離しながら音声波形をモデル化していることを示す。

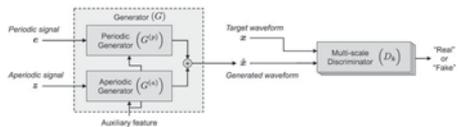


Fig. 1: Neural vocoder based on parallel model structure

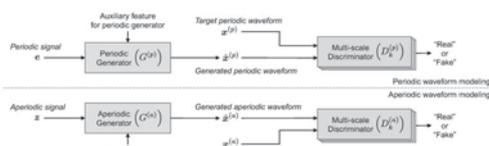


Fig. 2: Neural vocoder based on periodic and aperiodic decomposition

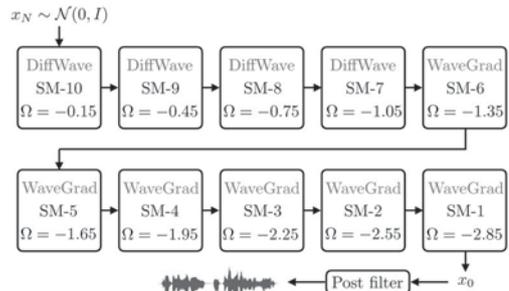
## 2-2-9

### 2-2-9 DiffWaveGrad: 拡散確率型ニューラル ボコーダの併用モデル

DiffWaveGrad: A combination model of diffusion probabilistic neural vocoders

○岡本拓磨<sup>1</sup>, 戸田智基<sup>2,1</sup>, 志賀芳則<sup>1</sup>, 河井恒<sup>1</sup>  
(<sup>1</sup>情報通信研究機構, <sup>2</sup>名古屋大学)

- ・拡散確率型ニューラルボコーダ: WaveGradとDiffWave
- ・時間領域の損失のみで学習可能なパラレル波形生成モデル
- ・拡散確率型ニューラルボコーダの改良→高品質化に成功
  - ・雑音レベルに対応したDiffWave
  - ・雑音レベルごとのサブモデリング→高音質化
  - ・高域強調ポストフィルタ
- ・DiffWaveGrad: WaveGradとDiffWaveの併用モデルの提案
  - ・雑音レベルの大きい推論の前半4回: DiffWave
  - ・雑音レベルの小さい推論の後半6回: WaveGrad



・結果: 高品質かつDiffWaveよりも高速な生成を実現

## 2-2-11

### 2-2-11 テキスト音声合成のためのポストフィルタ用 WaveNet ボコーダの学習条件に関する評価

Evaluation of training conditions of WaveNet vocoder as postfilter in text-to-speech.

☆安原和輝, Yi-Chiao Wu, Patrick Lumban Tobing (名大),  
松永悟行, 大谷大和(エーアイ), 戸田智基(名大)

- 従来型の TTS システムで生成される合成音声に対し、WaveNet ボコーダをポストフィルタとする枠組みが提案されている。
- ただし、論文中の合成音声は自然音声の韻律に基づくため、実際の TTS システムで生成される合成音声に対する性能評価は不明である
- 本研究では、実際の TTS システムで生成される合成音声に対して、ポストフィルタを適用し、WaveNet ボコーダの学習条件を変えて生成されるポストフィルタ音声への影響を調査する。

実験的評価により、WaveNet ボコーダをポストフィルタとする枠組みは、実際の TTS システムで生成される合成音声に対しても効果的であり、このような合成音声に合わせた学習を行うことで、品質をより向上させることを示す。

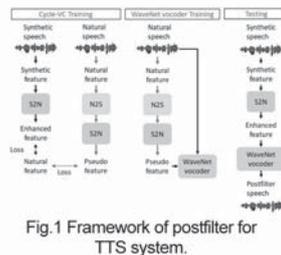


Fig. 1 Framework of postfilter for TTS system.

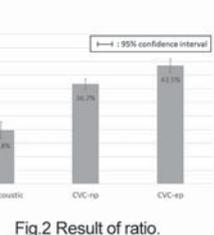


Fig. 2 Result of ratio.

### 2-2-12

#### 2-2-12 ハイブリッドオートエンコーダの強化学習による物理音声合成器の制御

Reinforcement Learning with Hybrid Auto-Encoder for Neural Network Based Automatic Physical Speech Synthesizer Control  
○篠崎隆宏, 柴田駿人, 張明鑫(東工大)

- ◆物理音声合成は、身体性を備えた音声発話ロボットや言語学習者への発声法教示、身体形状からの音声復元などに有用である。
- ◆物理音声合成器を駆動するニューラルネットを音声データのみを用いて教師なし学習する手法を提案した。
- ◆学習後のニューラルネットは任意の発話音声に対して動作し、物理音声合成器を駆動して再構成音声を得ることが可能である。

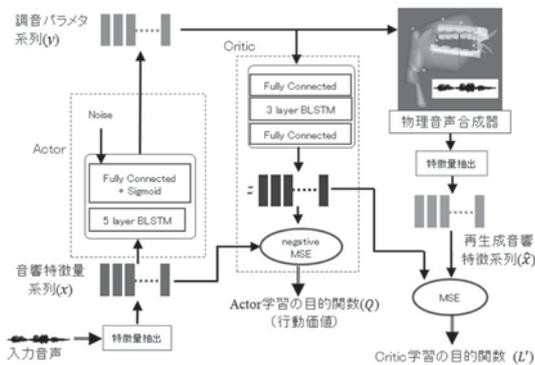


Fig.1: Hybrid auto-encoder based learning system for automatic physical speech synthesizer control. We use the deterministic policy gradient (DPG) for the learning.

### 2-2-13

#### 2-2-13 プライバシー保護のためのカスケード型音声加工法を用いた音声仮名化

Speech pseudonymization using cascaded voice modification for privacy protection

☆甲斐優人(都立大), 高道慎之介(東大), 塩田さやか, 貴家仁志(都立大)

- ◆本研究では、音声データを単独あるいは複数の信号処理法の組み合わせを用いて加工し、音声の仮名化を目的とする手法を提案する(Fig.1)。
- ◆EER と WER を用いて目的関数を定義し、目的関数の値が最小となるパラメータの最適化を行った。
- ◆音声の仮名化に単独よりも複数の加工法を組み合わせることで、すべての評価タスクで改善が見られた。複数の手法を組み合わせることでパラメータの探索空間を拡大し(Fig.2)、より適切なパラメータ選択を可能にしたことが影響している。

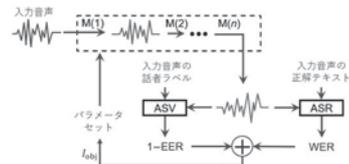


Fig. 1 : Proposed method to optimize the parameters

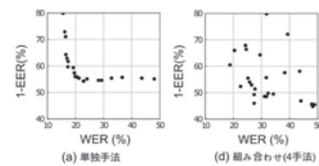


Fig. 2 : Parameter search space between single and four-method

### 2-4-1

#### 2-4-1 高齢者の音像定位能力と事象関連電位との関係

Relationship between event-related potentials and auditory lateralization ability in the elderly

☆白木萌子, △森田和元(中央大), △岩木直(産総研), 戸井武司(中央大)

高齢者の音像定位能力の把握を目的とした一連の実験により、特に右耳に早い両耳間時間差(ITD)を有した音で高齢者に誤定位が起こりやすいことが示唆された。本研究では、音像定位能力と事象関連電位との関係を調査するため、標準刺激に左右差のない音、逸脱刺激にITDを与えた音を用いてミスマッチ陰性電位(MMN)を測定した。

Fig. 1に高齢者の使用音源での音像定位実験の結果を示す。本実験においても、高齢者の右側の誤定位は確認された。Fig. 2に測定した脳波の波形例を示す。音が早く提示されたのと反対側の半球におけるMMNの音源提示後75-200msの区間平均電位が、負の方向に大きくなる傾向があることが分かった。特にITDによるMMNの差が認められ、ITDが0.4msの音源についてMMNが負の方向に十分大きくなっていった。

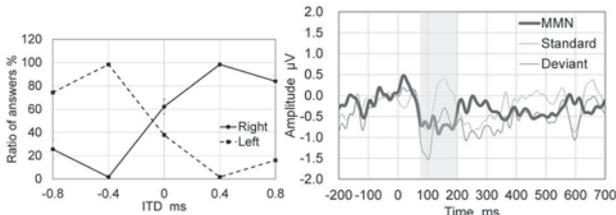


Fig.1 Ratio of answers for elderly participants

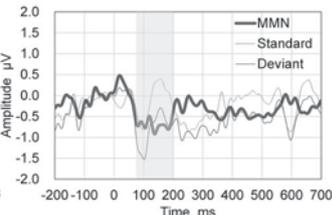


Fig.2 Example of ERP waveforms at C3 (ITD=0.4 ms for left ear preceding, younger participants)

### 2-4-2

#### 2-4-2 擬似片耳難聴状態の健聴者による水平面内音源定位

Sound source localization in the horizontal plane by normal-hearing listeners under simulated single-sided hearing loss

☆岡本しおり(千葉工大・院), 飯田一博(千葉工大・先進工)

- ◆片耳難聴者の音源定位能力の学習効果を検討する第1ステップとして、4名の健聴者に対して片耳難聴の初期状態を作り出し、水平面内の音源定位実験を行った。その結果、以下のことが示された。
- ◆4名中3名の被験者は片耳難聴状態では、音源方向に関わらず、健側に音像を知覚した。
- ◆1名は少なくとも左右方向については音源側に音像を知覚した。

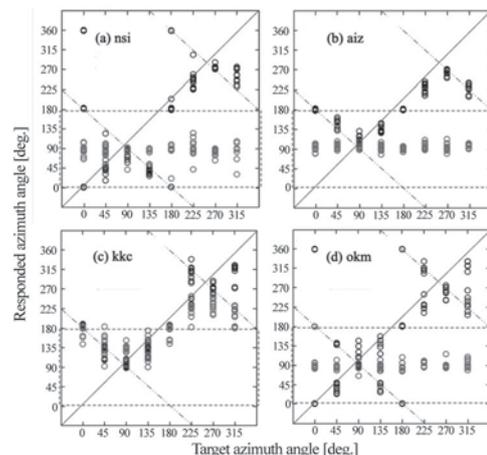


Fig.1: Directional localization responses to the stimuli in the horizontal plane. ○: normal hearing condition, ○: simulated single-sided hearing loss condition.

### 2-4-3

#### 2-4-3 合成バイノーラル信号の音像の分離知覚の検知限

Detection limit of sound images split by synthesized binaural signal

☆ 酒井翼, 森川大輔, モクタリパーハム, 平原達也 (富山県立大)

- ◆ 両耳間時間差 (ITD) と両耳間音圧差 (ILD) が同時に作用する条件下の音像の分離知覚の検知限を明らかにするために、頭部インパルス応答(HRIR)を用いて合成したバイノーラル信号の音像の分離知覚実験を行った。
- ◆ 検知限の2つの刺激音に畳み込んだ HRIR の角度差  $DL_{\theta}$  は、前方水平角0度から始まる条件 (F-cond) で約39度、後方水平角180度から始まる条件 (B-cond) で約43度であった。
- ◆ 検知限の HRIR を畳み込んだ2つの刺激音のそれぞれの ITD の差 ( $DL_{\theta}^{ITD}$ )、ILD の差 ( $DL_{\theta}^{ILD}$ ) を算出し、ITD または ILD のみ (ITD or ILD-cond) によって生じる音像の分離知覚の検知限と比較した。その結果、ITD は条件間の要因に有意差がなく、ILD の条件間の要因に有意差があった。

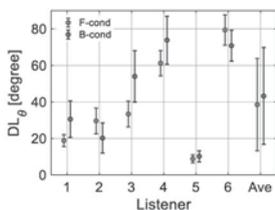


Fig1. 合成バイノーラル信号の  $DL_{\theta}$

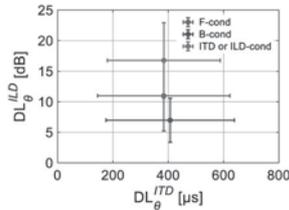


Fig2. 各条件の  $DL_{\theta}^{ITD}$  と  $DL_{\theta}^{ILD}$

### 2-4-5

#### 2-4-5 耳介形状数値データを用いた深層学習による個人の頭部伝達関数のノッチ周波数 bin のカテゴリ推定

Category estimation of notch frequency bins of individual transfer function based on deep learning using the anthropometry of the listener's pinnae

☆ 高橋仰一 (千葉工大・院), 飯田一博 (千葉工大・先進工)

- ◆ 個人化 HRTF を一般消費者が手軽に利用できるようにするため、深層学習を用いて耳介画像から個人化 HRTF を生成することを目指し、その第一段階として、耳介形状の数値データを用いて、正面方向 HRTF の N1N2 周波数 bin のカテゴリを推定した。
- ◆ N1N2 周波数 bin のカテゴリ分類は、周波数軸上で上昇角知覚に関する弁別閾に基づいて行ない、N1N2 それぞれ2カテゴリに分類された。
- ◆ K 分割交差検証法における的中率は、N1 では78.8%、N2 では66.4%であったが、訓練用耳介の水増しを行うことにより N1 では81.4%、N2 では72.3%となった。
- ◆ Leave Two Out 法における的中率は、N1 では71.8%、N2 では65.2%であったが、訓練用耳介の水増しを行うことにより N1 では76.1%、N2 では79.2%となった。

Table 1 Prediction accuracy for each evaluation method by DNN

evaluation method	N1 (%)	N2 (%)
K-Fold	78.83	66.36
K-Fold (data augmentation)	81.43	72.34
Leave Two Out	71.83	65.28
Leave Two Out (data augmentation)	76.06	79.17

### 2-4-4

#### 2-4-4 判別分析による正面と後方の個人の頭部伝達関数のノッチ周波数 bin のカテゴリ推定

Estimation of the category of notch frequency bins of the individual head-related transfer functions for front and rear directions using discriminant analysis

☆ 西山織絵 (千葉工大・院), 飯田一博 (千葉工大・先進工)

正面及び後方の HRTF の N1, N2 が含まれる周波数 bin のカテゴリを耳介形状から推定する方法を提案し、その推定精度を検証した。以下のことが明らかになった。

- 1) 正面方向において、76 耳介のうち 71 耳介の N1 は中心周波数 6.3 kHz から 8.4 kHz の 6 つの周波数 bin に分布した。同様に 72 耳介の N2 は中心周波数 9.0 kHz から 11.9 kHz の 6 つの周波数 bin に分布した。
- 2) 後方において、76 耳介のうち 70 耳介の N1 は中心周波数 7.5 kHz から 10.0 kHz の 6 つの周波数 bin に分布した。70 耳介の N2 は中心周波数 12.6 kHz から 16.9 kHz の 6 つの周波数 bin に分布した。
- 3) 上昇角知覚に関する N1, N2 周波数の弁別閾を用いると、6 つの N1, N2 周波数 bin はそれぞれ 2 つのカテゴリに分類できた。
- 4) ナイーブな耳介の正面方向の N1, N2 周波数 bin カテゴリの的中率はそれぞれ 83.1%、77.8% であった。後方では N1 は 74.3%、N2 は 77.1% であった。

Table 1 Percentage of correct classification of N1 and N2 frequency bin category for naive ears.

Front		Rear	
N1	N2	N1	N2
83.1%	77.8%	74.3%	77.1%

### 2-5-1

#### 2-5-1 新型コロナウイルスによるステイホーム中の音環境アンケート調査—建物構造別の居住者の意識調査結果—

Questionnaire survey on sound environment during stay at home by COVID-19 - Awareness survey results for residents by building structure -

○ 富田隆太 (日大・理工), △ 阿部今日子 (フリーランス)

- ◆ 新型コロナウイルスによるステイホーム中は、共同住宅全体の在宅率が高くなり在宅時間が長くなること等から、従来とは異なる騒音に暴露される可能性がある。このような環境は、共同住宅の騒音の負荷状態としては、最大に近いと推定される。
- ◆ アンケートは、インターネット調査により実施した。調査期間は、事前調査実施日が 2020 年 6 月 18 日～22 日、本調査実施日が 2020 年 6 月 22 日～26 日であった。
- ◆ 本報では、構造別の居住者意識について考察を行った。
- ◆ 問題意識の程度では、それぞれの構造で一般的に推定される性能に対して、居住者意識も概ね対応が取れている様子が見られた。これらは、長年にわたり建築技術者が遮音性能の向上に取り組んできた結果が居住者意識に反映されてきていると考えられる。
- ◆ 住まい方では、他住戸のために住まい方に配慮することが日本の文化として根付いていると考えられるが、住まい方に対する意識は、時代や生活スタイルの変化と共に変化していくことから、今後も継続して調査を行なっていく必要があるといえる。
- ◆ ステイホーム中という特殊な環境においても、多くの人々がトラブルに巻き込まれることなく過ごしていた中で、1 割前後の人がトラブルに巻き込まれていることは、見逃してはならない問題であろう。

### 2-5-2

#### 2-5-2 WEB 会議エリアのあるオープンプラン オフィスにおける音環境改善方法の検討 -実態調査と課題の提示-

Discussions of how to improve acoustic environment of open plan office with WEB conference area - Report of acoustic investigation of an open plan office and acoustic issues-

◎井上瑞紀(日建設計)

- ◆2020年のCOVID-19の流行以降、在宅勤務を導入する企業も増え、働き方が大きく変化した。本稿では具体的な音環境改善方法の検討にあたり、あるオフィスについてCOVID-19の影響により実際に執務者が感じている音環境の変化についてアンケート調査を行った結果と、現状のオープンプランオフィスの抱える課題について報告する。
- ◆COVID-19の流行によりWEB会議が急激に普及し、それに伴うオフィスの音環境の変化を執務者が感じていることが確認できた(Fig. 1, 2)。
- ◆WEB会議に関わる音環境について肉声とスピーカ音の気になる度合い、規模による妨害感の違いについて調査した(Fig. 3)。

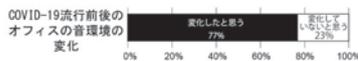


Fig. 1 The ratio of people who think acoustic environment has changed after pandemic



Fig. 2 Answers of the question "Are there any sound which make you disturbance more after COVID-19 pandemic?"

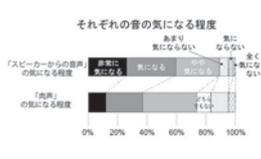


Fig. 3 Answers of the question "How much disturbance each sounds have?"

### 2-5-4

#### 2-5-4 飛沫防止用仕切り板の音響透過性能の向上 その1-有孔板を利用した方法-

Improvement of sound transmission performance of partition for droplet infection prevention, part 1: application of perforated plates.

◎杉江聡, 鈴木肇, 新田龍馬(小林理研)

- ◆新型コロナウイルス感染症の拡大防止に用いられる飛沫感染防止用の仕切り板の音響透過損失を低下させる手法を検討した。
- ◆Fig. 1のように、仕切り板の両面に有孔板を空気層を介して配置し、共鳴透過を生じさせることにより、遮音性能を低下させる。
- ◆片面のみに有孔板を配置した条件に比べ、両面に配置することによって、遮音性能が大きく低下する(Fig.2の実線と破線の比較)。
- ◆有孔板-仕切り板-有孔板の三層構造であり、2つの共鳴透過が生じると考えられるが、低い周波数に生じる共鳴透過の方が顕著に現れる。
- ◆その共鳴周波数は、仕切り板の厚さや密度には依存しない(Fig.2の破線と点線の比較)。

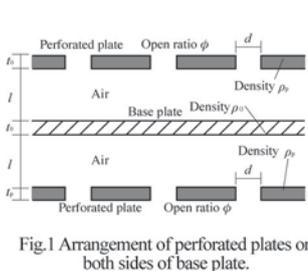


Fig.1 Arrangement of perforated plates on both sides of base plate.

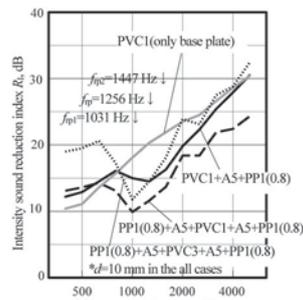


Fig.2 Comparison between measured results of perforated plates on both sides and single side.

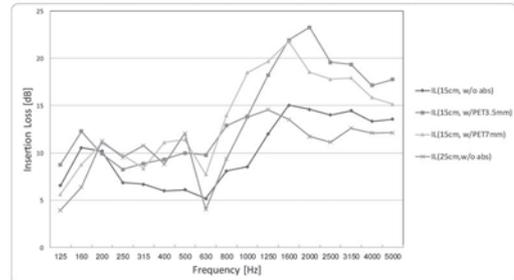
### 2-5-3

#### 2-5-3 ドアに対するプレナム構造の適用に関する 簡易実験\*

\*A simple mockup experiment on applicability of plenum structure to doors

○阪上公博, 奥園健, △村田遥, 田丸加奈子, △坪倉正佳, 草鹿みどり, 向江俊一(神戸大院・工)

- ◆プレナム構造は、開口を持つ自然換気機能を有する遮音構造として、蒸暑地域を中心に近年利用されており、研究事例も多い。一方、COVID-19の感染防止対策として、ドアを開放している商店等が見受けられるが、プレナム構造がドアに利用できれば、遮音上は有利であると考えられる。本報告では、簡易な実験により、プレナム構造のドアへの適用可能性の検討を試みた結果を述べる。
- ◆実験は一般的なオフィスが2室連結している間仕切りのドアを外した開口を利用して行い、開放時に対する挿入損失(IL)として評価した。その結果、プレナム構造は15~25cm幅の開口部を有するにも関わらず、予想以上の遮音性能が得られた。今後は、理論・実験両面からのさらなる考察を要するが、ドアへのプレナム構造の採用については、実用化への検討に値するものと考えられる。



Measured results of the insertion losses of the mock-up plenum structures. Opening widths: Blue:150mm (without absorber), Orange: 150mm (with 3.5mm absorber), Yellow: 150mm (with 7mm absorber), and Green; 250mm (without absorber).

### 2-5-5

#### 2-5-5 飛沫防止用仕切り板の音響透過性能の 向上 その2 -実用化の検討-

Improvement of sound transmission performance of partition for droplet infection prevention, part2: example of accrual application.

◎新田龍馬, 鈴木肇, 杉江聡(小林理研)

- ◆飛沫防止用の仕切り板に空気層を介して有孔板を配置し(Fig.1)、共鳴透過による遮音欠損が生じた時の透過側受音点の音圧レベルの変化を、回折音を含めた推定計算と模擬実験により検討した。
- ◆有限長障壁の回折音と、障壁の一部領域(有孔板取付部)を透過する透過音を独立に計算し、干渉を無視した受音点の音圧レベルを推定した。
- ◆音源(話者)と受音点(聴者)が仕切り板に近い場合に、仕切り板の両面の大部分を有孔板で覆うことで、仕切り板単体の場合と比べ、受音点の音圧レベルが透過損失のディップとなる周波数で約5dB向上した。
- ◆有孔板の面積を縮小させると、回折音に対する透過音の割合が小さくなり、受音点の音圧レベルの向上はわずかとなった。
- ◆アクリル製の仕切り板と、塩ビ製の有孔板を用いた実験(Fig.2)を行い、推定計算結果と同様の傾向が示された。

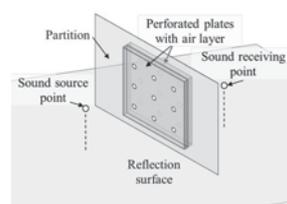


Fig.1 Geometric positions of partition, sound source and receiving position.

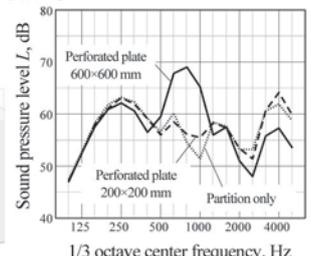


Fig.2 Measured results at sound receiving point for various area of perforated plate.

## 2-5-6

### 2-5-6 アクリルパーティションへのMPPの応用

Application of MPPs to an acrylic table partition

☆草鹿みどり, 阪上公博, 奥園健(神戸大・院), 山口大地, 城戸重幸(NC産業)

- ◆本研究では、COVID-19の感染拡大防止策として飲食店等に設置されるアクリルパーティションに、微細穿孔板(Micro-Perforated Panel, MPP)を応用させたものを提案し、吸音特性の検討を行った。
- ◆試験体はアクリル板の両側に、空気層を挟んでMPPを設け、3重としたパーティションである。試験体は上側と両側面の3辺が開放となるため、3辺をガムテープまたはアクリルの蓋で塞いだ試験体も作成し、3つの異なる条件で残響室法吸音率を測定した。
- ◆試験体1個当たりの等価吸音面積で比較したところ、ガムテープで塞いだ試験体、アクリルの蓋で塞いだ試験体、3辺開放の試験体の順に高いピークを示した。開放部分の3辺を塞ぐことで、音波が側面から入射または漏れることを防ぎ、吸音性能が向上したと考えられる。

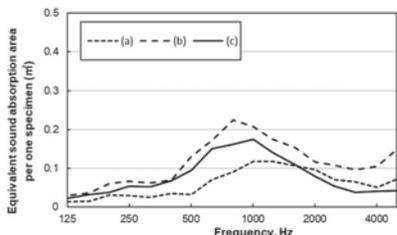


Fig. 1: The comparison of equivalent sound absorption area per one specimen in three different conditions, (a): specimens with three edges open, (b): specimens with three edges closed by gum tapes, (c): specimens with three edges closed by acrylic covers.

## 2-6-2

浴室における人の有無などの検知  
—音響特性の計測による室内状態の検知(IV)—

Detection of person in a pre-fabricated bath unit.

— Detection method of room conditions by measuring acoustic characteristics (IV) —

☆川部 良隆, 和田 有司, 中村 健太郎(東工大・研究院)

- ◆ 独居高齢者の増加により、カメラなどの遠隔監視装置を用いた見守りサポートの需要が増加している。しかし、浴室やトイレにカメラを設置することは、プライバシーの観点の問題が大きい。
- ◆ これまで、室内の音響周波数特性を測定することにより人の有無や人数、室温の推定ができることを示してきた。
- ◆ 本報告では、特にカメラを利用しにくい浴室を想定して人の有無や姿勢を検知することを試みた。浴室は比較的狭い閉じた空間なので音響モードが明確に現れるほか、インターホン機能のあるコントローラが設置されている場合が多く、本手法を適用しやすいと考えた。
- ◆ 人の有無による室内音響特性の変化は、ピーク周波数の移動やゲインの大きさなどに現れる。
- ◆ Fig. 1のように、室内を模したアクリルケース(1/4縮小)内にスピーカからスイープ音を出しマイクロホンで受信し、ケース内の物体の有無や姿勢変化による周波数特性の変化を調べた。
- ◆ 角柱や美術用人形の有無が推定できた。
- ◆ 角柱を横に置いた場合、縦に置いた場合と底面の境界条件変化から周波数特性が大きく変化する(Fig. 2)。これから姿勢の推定も可能であると考えられる。

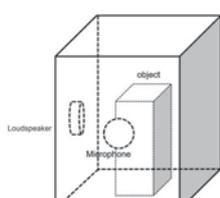


Fig. 1 Experimental setup.

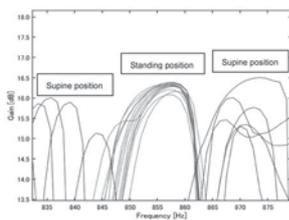


Fig. 2 Frequency characteristics

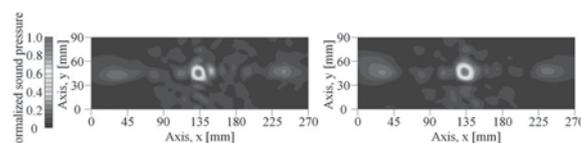
## 2-6-1

### 2-6-1 時間反転波による空中超音波フェーズアレイの音波集束の検討

Examination of sound wave focusing of airborne ultrasound phased array by time reversal wave

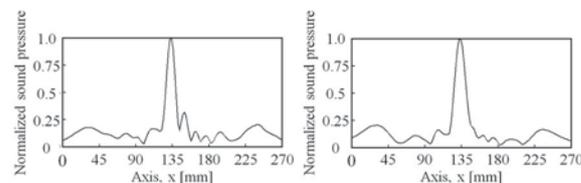
☆清水鏡介, 大隅歩, 伊藤洋一(日大理工)

- ◆ 空中超音波フェーズアレイ(Airborne Ultrasound Phased Array : AUPA)と弾性波源走査法を組み合わせた非破壊検査方法について研究を進めている。
- ◆ 本報告では、時間反転波を用いたAUPAの音波集束について検討を行った。
- ◆ 結果より、時間反転波による音波の集束について確認した。



(a) Phase control method (b) Time reversal method

Fig. 1 Normalized sound wave focusing distribution.



(a) Phase control method (b) Time reversal method

Fig. 2 Normalized sound wave focusing distribution (y = 46mm).

## 2-6-3

### 2-6-3 音波照射加振による非接触浅層地中探査法の基礎検討

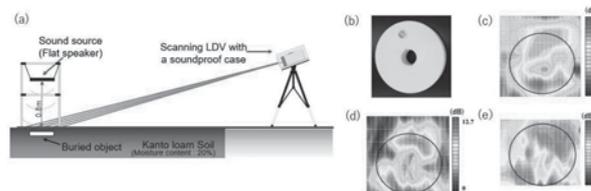
—関東ローム土を用いた場合—

Basic study of noncontact shallow underground exploration method by acoustic irradiation induced vibration, -When using Kanto loam soil-

○杉本恒美, 杉本和子, 中川 裕, 上地 樹(桐蔭横浜大・院),

坂本憲昭, 高木 均(桐蔭横浜大・院), △歌川紀之, 黒田千歳(佐藤工業)

- ◆ 比較的浅い地中を探査する手法としてはレーザ探査、磁気探査、電気探査および弾性波探査といった方法が主に用いられている。
- ◆ しかしながら、どの手法も計測用機材を地表面に接触もしくは近づける必要があり、遠隔から非接触で探査を行うことは困難であった。
- ◆ 一方で、音波照射加振を用いた非接触音響探査法では、小型音源を搭載したUAVを用いれば10 m以上の離隔があったとしても、非接触で浅層地中探査が実現できる可能性があるため、基礎検討を行っている。
- ◆ 今回は関東ローム土を用いた基礎検討を行い、レーザドップラ振動計から20 m程度離れた位置からでも埋設物の形状を反映したドーナツ状の振動が計測できることを確認した。



(a) Experimental Setup. (b) Photograph of a buried object (Plastic pole base (dia; 300mm, depth; 50mm), Vibration energy ratio distribution (50-200 Hz, Kanto loam soil, water content 20%) (c) 10 m, (d) 20 m, (e) 30 m.

### 2-6-4

#### アニュアラレイによる後方散乱特性解析の臨床応用に向けた安定性評価

Stability evaluation for clinical application of backscattering characteristic analysis by annular array

☆西東若菜(千葉大・院融合), 大村眞明(富山大・工/千葉大・CFME),

△Jeffrey A. Ketterling(Riverside Research), 山口匡(千葉大・CFME)

- ◆中心周波数 20 MHz のアニュアラレイで散乱径径が 5, 10, 20, 40  $\mu\text{m}$  の 4 種のファントムを計測し, それぞれのエコー信号を評価対象および参照媒質として用いて, 後方散乱係数 (BSC) を評価した。また, 安定性評価として, 単一凹面振動子との比較を行った。
- ◆アニュアラレイと単一振動子の比較では, 周波数依存性の傾向は同じであるが, BSC の平均値と分散に違いが確認された。参照媒質が粒径の大きいファントムでは, 特に高周波数帯において単一振動子の分散が大きい結果となった (Fig. 1)。
- ◆ビームフォーミング (Synthetic Focusing) を適用したアニュアラレイの送受信音場特性の利点を活かした広い空間領域での組織性状評価の可能性が示唆された。今後は, 臨床応用に向けたリアルタイム性の向上を踏まえたシステム構築を行う。

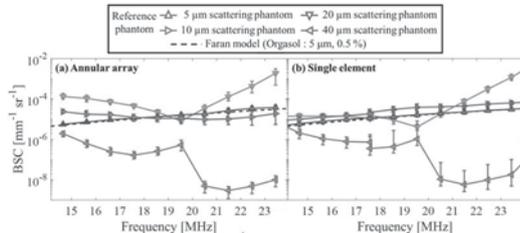


Fig.1: BSC curves of the 5  $\mu\text{m}$  scattering phantom.

### 2-6-6

#### 2-6-6 心筋線維の高分解能観察を目指した超音波送受信系の伝達関数の深さ依存性を考慮したフィルタ設計

Filter design considering depth dependence of transfer function of ultrasound measurement system for high resolution observation

☆川又健太, 森 翔平, 荒川元孝, 金井 浩(東北大)

- ◆心筋線維の走向の高分解能観察のために, 受信信号を広帯域化するためのフィルタの設計・適用を行っている。本報告では, その設計過程において, 超音波送受信系の伝達関数  $H(\omega; d)$  の深さ依存性を考慮することの重要性について検討を行った。
- ◆まず, 水中に設置したワイヤの深さ  $d$  を 5~55 mm まで変化させて取得した受信 RF 信号より, 深さごとの伝達関数を推定した。推定した  $H(\omega; d)$  を用いて, 逆フィルタを設計し, 各深さに設置したワイヤからの信号に対し, 各々フィルタの適用を行った。
- ◆Fig. 1(a)に, 深さ依存性を考慮したフィルタの適用結果, Fig. 1(b)には, 深さ依存性を考慮しないフィルタの適用結果を示す。これらの比較により, Fig. 1(a)の方が信号強度が強く, 空間分解能が向上していることが分かる。したがって, フィルタ設計の際に, 伝達関数の深さ依存性を考慮することは重要である。

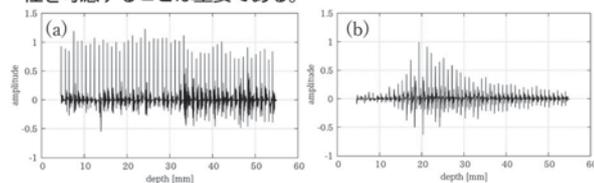


Fig.1: Results of applying inversed filter

(a) considering / (b) without considering depth dependency of transfer

### 2-6-5

#### 2-6-5 ラット肝臓の組織構造と後方散乱特性の関係性の基礎検討

Basic study to evaluate relationship between backscatter characteristics and tissue structure of rat liver

☆八木風音(千葉大・工), 西東若菜(千葉大・院融合),

大村眞明(富山大・工, 千葉大・CFME), 平田慎之介, 山口匡(千葉大・CFME)

- ◆複数の超音波センサを用いて 2~35 MHz の広帯域においてファントムおよびラット肝臓のエコー信号を収集し, 正常組織および脂肪や線維造が混在したラット肝臓の後方散乱係数解析と散乱体サイズ推定を行うことで, 組織構造と後方散乱特性の関係性について考察した。
- ◆どの周波数帯においても, 正常肝 (Control) に比して脂肪肝 (Fatty I, II), 線維肝 (Fibrosis) の BSC が大きい傾向が確認された。
- ◆各肝臓の散乱体サイズ ESD (Fig.1) 推定において, structure factor model (SFM) を用いた場合には, Control に比して Fatty や Fibrosis の ESD が顕著に大きく評価された。
- ◆これは, SFM が構造を有する散乱体を想定したモデルであり, Control では小さい散乱体, 脂肪化の進んだ Fatty II では大きい散乱体が均質に存在するものとして評価され, Fatty I と Fibrosis では様々な大きさの散乱体が混在した状態として評価されたためと想定される。

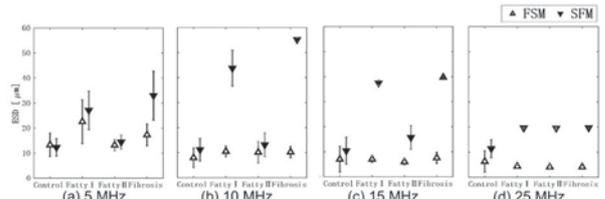


Fig.1: Estimated scatterer size of rat livers.

### 2-6-7

#### 2-6-7 (招待講演)利用技術から考える DML の可能性

(Invited Talk) Distributed Mode Loudspeaker (DML) potential with application technologies.

○河原 一彦(九州大・芸工),

- ◆Distributed Mode Loudspeaker (DML)は, 1990 年代後半に英国 New Transducers Limited (NXT 社)により提案されたパネル型スピーカの実装方式の一つである。
- ◆本稿では DML の放射の性質について着目し, DML の利用技術について概説する。

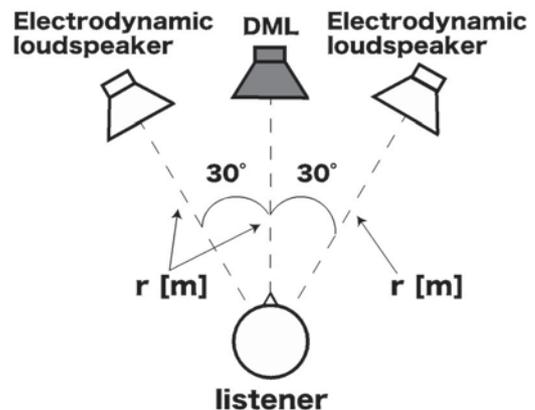


Fig.1: Example of a stereo sound reproduction system with a complementary distributed mode loudspeaker

### 2-6-8

#### 2-6-8 (招待講演)ヴァイオリンのモード振動における実験と数値シミュレーションの比較検証

Modal analysis of violin by experiment and numerical simulation, by YOKOYAMA, Masao (Meisei University).

○横山 真男

- ◆ オールドイタリアンのヴァイオリンの研究として、本体のモード解析について古典的なクラドニ法による解析と、数値シミュレーションによる比較解析について紹介する。
- ◆ マイクロCTスキャンによる詳細な3次元形状データを取得し振動と音場の連成数値シミュレーションを行っている。



Fig.1 Vibration mode by Chladni method (right: back plate, mode 2)



(a) A0 275.4Hz (b) CBR 425.8Hz  
Fig.2 Displacement of violin in mode A0 and CBR

### 2-6-10

#### 2-6-10 騒音低減のための構造設計法とその応用

Design method for reduction of noise and vibration and its application

○吉村卓也, 洪佳駿(都立大)

- ◆ 振動騒音の効果的な低減のためには、適切な構造変更を行うことが重要であり、そのような構造変更箇所を探る方法として感度解析がある。
- ◆ 感度解析とは、パラメータを微小に変化させたときの応答量の変化を微分係数で表したものであり、本稿では質量付加感度、剛性付加感度を用いている。
- ◆ 例えば自動車ボディの騒音は主にパネル等薄板構造物からの放射音为主体となっており、このような振動騒音は曲げ振動が主要因である。
- ◆ 平板の振動応答に対する感度解析を実施し、①~⑩の候補点の中から適切な構造変更箇所を探索した。効果的な剛性補強箇所を選定し小さい板を付加固定した。これによりピーク近傍の振動パワーを効果的に低減することができた。

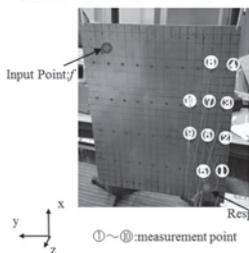


Fig. 1 Objective plate structure

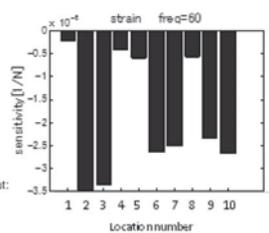


Fig. 2 Stiffness sensitivity around the resonance

### 2-6-9

#### 2-6-9 分布モードスピーカの立体構造化による新展開

New development by three-dimensional structure of distributed mode loudspeaker.

○大久保 寛(都立大), 黒澤 将(都立大)

- ◆ 近年、平板に分布させた振動モードを陽に利用するスピーカ(分布モードスピーカ(DML))が提案されている。
- ◆ 著者らは、従来のDMLの問題点を解決するため、シェル型構造を有する板面を振動板として使用する「シェル構造型DML」を提案した。
- ◆ シェル構造型DMLは、従来の平板に比べて構造的な自由度が高く、板面の形状をパラメータにすることで、原理的には最大寸法を制御しつつ、スピーカの周波数特性を最適化することも可能となる。
- ◆ 本稿では、従来の平板構造であったDMLをシェル構造によって3次元に立体構造化する手法を述べ、複数の種類のシェル構造型DMLの比較を行い、形状に対する振動や音響放射の特徴について報告する。

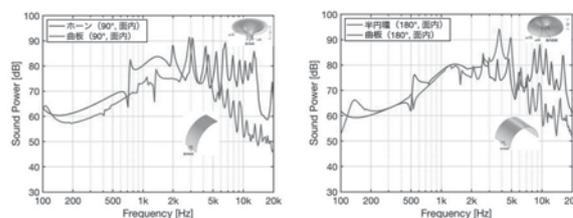


Fig.1: シェル構造型DML(部分円筒型DML及びホーン型・半円環型)の音響放射パワー

### 2-6-11

#### 2-6-11 (招待講演)光学的音響計測を用いたスピーカ放射音の可視化計測と解析

Measurement, Visualization and Analysis of Sound Radiation using Optical Method

○及川靖広(早大理工)

- ◆ 近年の光学的音響計測技術により、可聴音についても空間的な伝搬が観測しやすくなり音場の直感的な理解が容易になってきた。その計測結果の解析方法についても検討が進んでいる。本稿では、スピーカ放射音を対象として行なった光学的音響計測による可視化結果とその解析について紹介する。
- ◆ 光の音による位相変化を計測するには、干渉法を用いることが多い。我々はレーザドプラ振動計(LDV)を用いたシステム、偏光高速度干渉計を用いたシステムを主に使用している。それぞれに長所、短所があるので、計測対象、計測目的に応じて使い分けると良い。その計測結果は音場の2次元イメージングとなることに注意が必要である。
- ◆ 空間スペクトルを解析することにより、音の振幅に加え、周波数・含まれる波面の向きを同時に観測することが可能である。

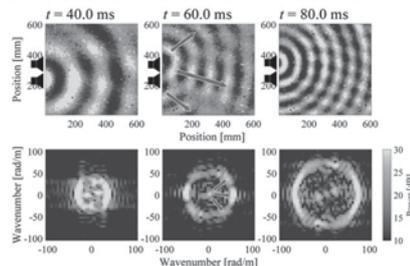


Fig.1 Sound field and spatial spectrum for 2-way loudspeaker.

2-6-12

2-6-12 サイクルオクターブウェーブレット変換を用いた分布モードスピーカ音響振動解析

Vibro-acoustic analysis of Distributed Mode Loudspeaker using cycle-octave wavelet transform

☆黒澤 将, 大久保 寛 (都立大)

- ◆ 平板上に屈曲波モード発生させることによって音を放射する「分布モードスピーカ(DML)」が90年代より盛んに研究開発されている。
- ◆ 当研究グループはシェル形状を有する板をDMLの振動板として使用する「シェル構造型DML」を提案した。シェル構造型DMLは、面内方向に進む波によって共振現象(面内波モード)が生じ、その周波数において強い音響エネルギーを放射する。
- ◆ 音響パワー周波数特性などの定常特性の評価ではDMLの重要な特徴の一つである振動板の残響応答についての特性を捉えることは難しい。
- ◆ そこで、DMLの新しい評価指標としてサイクルオクターブウェーブレット変換を導入し、有限要素法・境界要素法を用いた音響振動解析と組み合わせて、平板と曲板の残響の定量化を図った。
- ◆ ウェーブレット応答結果の比較により、平板に生じる屈曲波モードと曲板に生じる面内波モードの時間-周波数特性の違いを明らかにした。

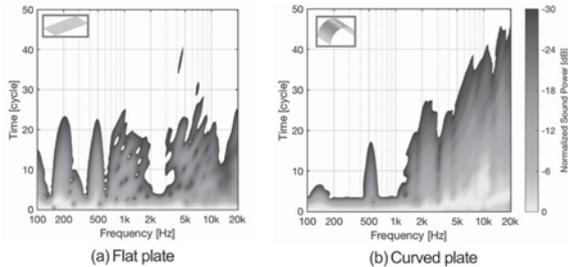


Fig.1: Results of cycle-octave wavelet transform

2-7-2

2-7-2 たわみ振動板型空中超音波音源から放射された音波のL字型反射板による集束

Focusing of aerial sound waves emitted from the sound source using transverse vibrating plate attached L-shaped type reflective plates.

☆田名部美穂, 浅見拓哉, 三浦 光 (日大・理工)

- ◆ 空中に強力な超音波を放射する音源の一つに、Fig. 1に示すような凸端駆動モード矩形たわみ振動板を用いた音源がある。
- ◆ 本稿では、Fig. 2に示すようなL字形の反射板を設置して、放射音波を一方に集束させ、大容量の音波を放射する音波放射機構を考え、その集束方法について検討を行った。
- ◆ その結果、L字型反射板の長さは音波の半波長の整数倍とし、下側の反射板のみ音波の半波長分長くすることで、音波が一方に集束し、高い音圧が得られることを明らかにした。

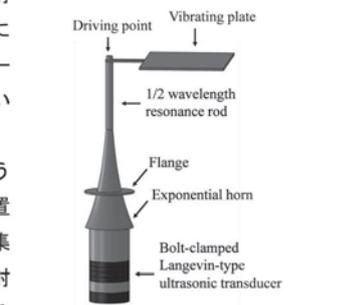


Fig. 1. Outline of an ultrasonic source.

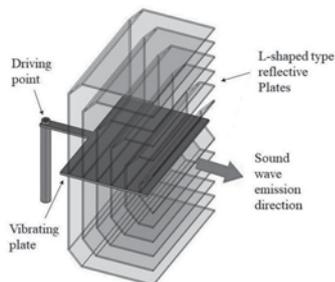


Fig. 2. Sound wave radiation mechanism.

2-7-1

2-7-1 円筒形振動板型超音波音源による煙霧質凝集の基礎検討

Basic study of aerosol coagulation by ultrasonic sound source using a cylinder type vibrating plate with rigid wall

☆宝田祐介, 浅見拓哉, 三浦 光(日大・理工)

- ◆ 本研究では、空中強力超音波による凝集率の向上を目的として、定在波音場を用いた煙霧質の凝集を行っている。本稿では、新しい形式の円筒形たわみ振動板型空中超音波音源を用いて凝集を行い、異なる入力電力における微粒子の凝集率について検討を行った。
- ◆ Fig. 1は、本検討で行った実験結果である。微粒子の質量濃度は超音波照射直後に急激に減少し、しばらくすると一定値に落ち着くことが分かった。また、凝集率は音源の入力電力が10Wで約45%、50Wで約85%となり、入力電力が大きくなると、向上することが分かった。

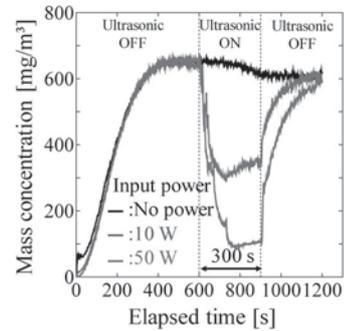


Fig. 1 Relationship between mass concentration and elapsed time.

2-7-3

2-7-3 ヘリカルスリット付き伝送棒を用いた複合振動源の振動特性

Vibration characteristics of complex vibration source using transmission rod with helical slits

☆宮田義大, 浅見拓哉, 三浦 光(日大・理工)

- ◆ 筆者らは、超音波接合の接合強度を向上する方法について検討している。従来の超音波接合は直線振動のみによる接合が多く、接合強度が低いなどの問題があった。そのため、筆者らは複合振動を生じる振動源の開発を行っている。
- ◆ 本稿では複合振動を発生させるために、Fig. 1に示すようなヘリカルスリットを施した伝送棒を用いた複合振動源の振動特性について検討した。

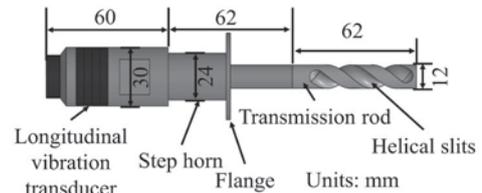


Fig. 1. Outline of ultrasonic vibration source.

- ◆ スリットの深さを变化させた場合における伝送棒先端での縦ねじり振動共振周波数時の振幅比(=ねじり振動振幅/縦振動振幅)の変化を検討した。
- ◆ その結果、縦振動共振周波数時の振幅比は、ヘリカルスリットが深くなるほど増加した。一方、ねじり振動共振周波数時の振幅比は、ヘリカルスリットが深くなるほど減少することが分かった。

### 2-7-4

#### 2-7-4 細長い接合用チップを用いたダンベル型縦-ねじり複合振動源の開発

Development of dumbbell-shaped longitudinal and torsional complex vibration using an elongated welding tip

☆安藤 駿, 浅見拓哉, 三浦 光(日大・理工)

- ◆細長い接合チップを取り付けたダンベル型超音波複合振動源を作成し、接合チップ部に生じる縦及びねじり振動方向に平行なたわみ振動の振動変位について検討し、チップ先端で振幅の最大値が得られる条件を明らかにする。
- ◆矩形断面の一樣棒を接合チップとして、超音波振動源の縦振動共振周波数である 29.4 kHz と、ねじり振動の共振周波数である 18.6 kHz でたわみ振動の共振をもつ接合チップの寸法を有限要素法解析にて求めた。その結果、長さ 100mm の場合、断面は縦振動方向 4.6mm、ねじり振動方向 4.2 mm を得た。
- ◆解析にて求めた寸法の接合チップを製作し、各共振周波数にて駆動させた。Fig.1 はそのときのチップ部の振動変位である。図より、両方向の振幅とも先端（100 mm の位置）にて振幅の最大値が得られていることがわかった。

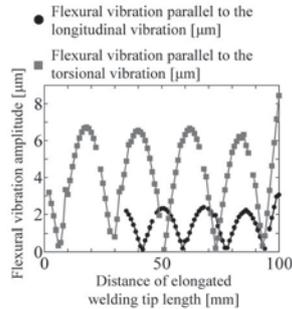


Fig. 1 Vibration distribution of elongated welding tip

### 2-7-6

#### 2-7-6 音場による光路長変調を利用した 100 kHz 空中超音波の測定

Measurement of 100-kHz airborne ultrasound through modulation of optical path length

☆佐藤 純平, 和田 有司, 中村健太郎(東工大)

- ◆空中超音波による manipulation や haptic 技術の開発が進み、より高い分解能のために従来(40 kHz)よりも高周波数の利用が期待されている。このため、100 kHz 程度の空中超音波の測定技術が必要とされている。
- ◆本報告では、音場による光学屈折率の変調効果を利用し、干渉計の光路長変調を起こすことで 100 kHz の空中超音波を測定することを検討した。
- ◆レーザドップラ振動系の測定光を 2 次元状に展開することで 2 次元受音面とし、鋭い指向特性をめざした。
- ◆Fig. 1 のように 52 mm × 45 mm の範囲にレーザ光路を展開したところ、Fig. 2 のように理論通りの 5~6° の半値全幅の鋭い指向特性が得られた。
- ◆垂直面指向性のサイドローレベルが高いのは床面からの反射波の影響と考えられる。

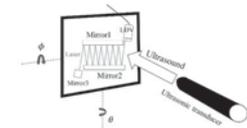


Fig. 1 Setup for directivity generation.

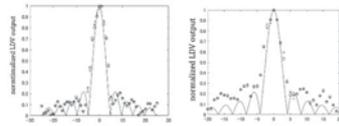


Fig. 2 Horizontal directivity (left) and vertical directivity (right).

### 2-7-5

#### 2-7-5 超音波振動が神経突起の成長に与える影響

The effects of ultrasound vibration on neurite growth

☆丸山悠輝, 藤原滉二, 小山大介(同志社大)

- ◆神経機能回復のために、神経幹細胞から分化、誘導した細胞を用いる治療法が検討されている。
- ◆細胞培養ディッシュ底面に超音波振動を励振し、神経突起の成長方向の制御を試みた。
- ◆PC12 細胞に神経成長因子を添加し、約 20 時間培養した後超音波振動を 72 時間励振させ、神経突起の成長方向および長さを計測した。
- ◆神経突起の成長方向は、Control が各角度に一樣に分布したのに対し、超音波振動を励起した場合は 90° 方向（接線方向）に分布したことから、超音波振動によって神経突起の成長方向の制御が可能であることが示唆された (Fig. 1)。
- ◆神経突起の長さは、超音波振動および培養環境の温度上昇により成長が促進されることが確認された。

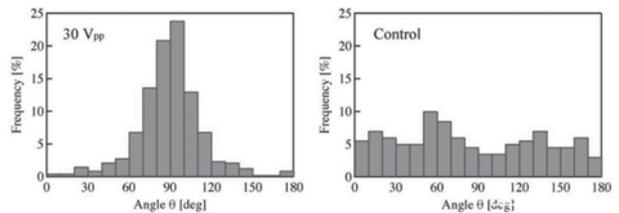


Fig. 1 Neurite growth directions on the culture dish without and with ultrasound excitations.

### 2-7-7

#### 2-7-7 強力空中集束超音波による物体表面温度の数値シミュレーション

Numerical simulation in temperature on material surface under high-intensity aerial focus ultrasonic wave

☆菱沼真志, 大隅歩, 伊藤洋一(日大・理工)

- ◆強力超音波照射によるエネルギーを利用した技術に、HIFU による癌治療やサーモニック法を利用した非破壊検査がある。
- ◆これらの応用技術の一つとして、強力空中集束超音波 (HIAFU) を利用した非接触加熱技術がある。一方、この技術を実用化するためには、音波照射時間と温度上昇、照射範囲と温度上昇分布等の関係を詳細に知ることが不可欠であり、そのための実験やシミュレーションによる検証が必要となる。
- ◆本報では、HIAFU 照射によるシリコーンゴムの表面温度上昇の数値シミュレーションの基礎検討を行い、実測との比較を行った。

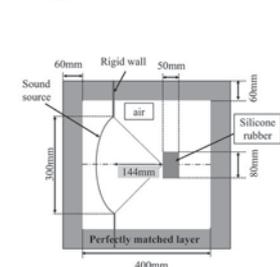


Fig. 1 Schematic view of analysis model.

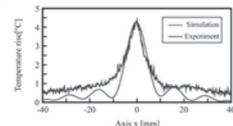


Fig.2 Temperature rise distribution along silicone rubber surface.

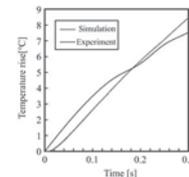


Fig.3 Temperature rise

2-7-8

2-7-8 強力空中超音波を用いた弾性波源走査法  
励起による Lamb 波の伝搬シミュレーション

Numerical simulation of Lamb wave propagation  
by wave source scanning method using aerial ultrasonic wave

○山田健太, 大隅歩, 伊藤洋一(日大・理工)

- ◆強力空中超音波による弾性波源走査法を利用した材料内部の欠陥を検出する方法について基礎検討を行っている。
- ◆本報告では, 提案手法における空中超音波の波源径が欠陥可視化に与える影響について, 有限要素法による数値シミュレーションと実測の比較検証を行った。
- ◆その結果, 空中超音波の波源径 D が, 欠陥可視化に影響を与えていることを確認した。

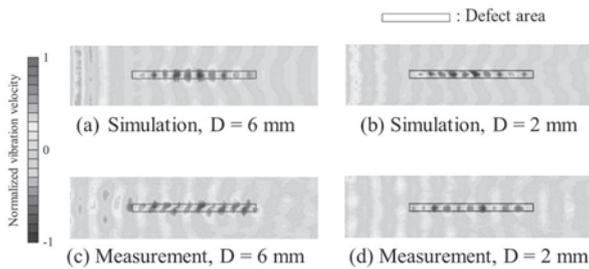


Fig.1: Lamb wave propagation (200 kHz)

2-7-10

2-7-10 金属円筒の第一次軸対称振動モードの  
励振方法の検討

Examination of excitation method for the first axisymmetric vibration mode  
of metal cylinder.

☆村井海斗, 孔徳卿(室工大・院), 田村英樹(東北工大), 青柳学(室工大・院)

- ◆Fig.1 に第一次軸対称振動モード (L(0,1) モード) を示す。金属円筒の両端部に圧電円筒振動子を設け, 径方向にひずみを与えることで励振する。
- ◆Fig.2 に示す試作した振動子の解析モデルを用いて, 有限要素法により進行波励振の確認を行った。金属円筒を共振長より  $\lambda/4$  だけ長くし, 振動子を  $\pi/2$  位相差駆動によって進行波の励振が確認された。
- ◆Fig.3 に 2 電源位相差駆動時に金属円筒表面の径方向変位の測定結果と有限要素解析結果を示す。最大振幅が規格化されている。定在波比が Position = 30 mm 付近において約 0.8, Position = 105 mm 付近で約 0.55 の結果が得られた。物体搬送への応用には, 円筒を強励振する方法が課題である。

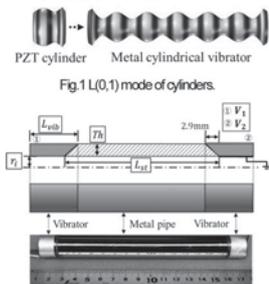


Fig.2 Model of cylinder.

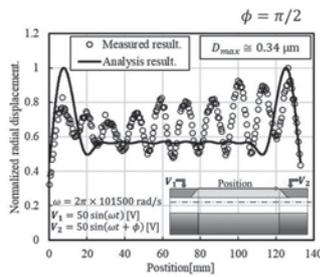


Fig.3 Radial vibration displacement of pipeline.

2-7-9

2-7-9 圧電円盤振動子の厚み振動モードの  
有限要素法解析

Finite element method analysis of thickness vibration mode  
of piezoelectric disk oscillator.

☆花田 徳, 青柳 学, 孔 徳卿(室蘭工大・院)

- ◆本報告では, 水と空気の放射力差を用いて駆動する液中推進システムをもとに, 有限要素法解析を用いて 2 つの材料 C-204 と C-213 を比較した結果について報告する。
- ◆固有周波数解析により得られる厚み振動モードとその周波数を Table 1 に示す。また, アドミタンス特性の測定結果を Fig.1 に示す。
- ◆C-204 では 2.0656MHz, C-213 では 2.0153MHz において, 周波数領域解析を用いて空気負荷, 水負荷を加えたときの放射面における振動振幅の Z および R 成分と音響放射力の計算結果を Table 2 に示す。これより, 1V<sub>0</sub> の同電圧の条件下では振幅および音響放射力は C-213 の方が大きくなることが確認できた。

Table 1 Vibration modes and frequency (MHz) analysis results.

	Mode 1	Mode 2	Mode 3
C-204	1.9428	2.0656	2.1467
C-213	2.0153	2.1438	2.2232

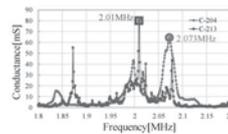


Fig.1 Admittance characteristics in air.

Table 2 Performance of C-204 and C-213.

	C-204	C-213	
空気側振幅	Z成分 [mm]	0.210	0.260
	R成分 [mm]	0.061	0.076
水側振幅	Z成分 [mm]	0.205	0.262
	R成分 [mm]	0.053	0.078
空気音響放射力 [mN]	$2.01 \times 10^4$	$2.73 \times 10^4$	
水音響放射力 [mN]	0.118	0.151	

2-7-11

2-7-11 超音波曝露の許容値に関する検討

Study on Tolerance of Ultrasound Exposure

○星貴之(ピクシーダステクノロジーズ株式会社)

- ◆近年, パラメトリックスピーカの用途展開や超音波フェーズドアレイの登場により強力な超音波がより身近なものになってきている。
- ◆従来の安全基準は 1900 年代半ばの少数の論文に依拠するものであり, またそれらの論文は 20 kHz 周辺の低い周波数帯に関するものが多い。
- ◆近年応用が広がっている周波数帯は 40 kHz 周辺の高い周波数帯であるため, 改めて検討し直す必要がある。
- ◆筆者は有志(大学・研究所の研究者, 関連企業の担当者)の勉強会を立ち上げ, 安全基準について議論した。
- ◆140 dB SPL 未満の 40 kHz 超音波が危険であるというエビデンスは現在のところ存在しないことを指摘した。また, 音圧レベルによらず違和感を覚えた際には使用を中止することを推奨した。

Effect	Body part	Threshold [dB SPL]
Tactile	Heat	Skin surface: 188.7 (estimated)
		Skin cleft: 140
	Vibration	Palm: 136.4 - 141.5
	Airflow	Face: 163.0 (lower jaw)
Audio	Hearing	Ear: 133.3 - 140.2
		Face: 166.3 (cheek beside nose)
	TTS	Ear: 148

[Note] Only valid for around 40-kHz ultrasound

Fig.1: Summary of effects of ultrasound on human body.

## 2-7-12

### 2-7-12 空中超音波の定在波音場を用いた液滴の操作

Manipulation of droplets using standing wave field of ultrasound in air

☆谷智司(愛工大), 畑中信一(電通大),  
△豊田昌弘(本多電子), 小塚晃透(愛工大)

- ◆小型超音波スピーカ 36 個を凹面形状の土台に配置した超音波ユニットを 2 台向かい合わせて、集束する定在波音場を形成する。超音波スピーカに印加する電圧を変化させて、音圧を変化し、様々な種類の液滴を音圧の節に捕捉する。そして、超音波の位相を変化させて、捕捉した液滴を移動させる操作を行い、評価を行う。
- ◆Fig. 1 に実験で使用した超音波ユニットの(a)全体、(b)背面、(c)振動子面の写真を示す。また、Fig. 2 に、30 V<sub>pp</sub> の電圧を印加して、形成した超音波の定在波音場中の音圧の節に、水滴が捕捉されている様子の(a)全体、(b)液滴の拡大図、(c)斜め方向の写真を示す。

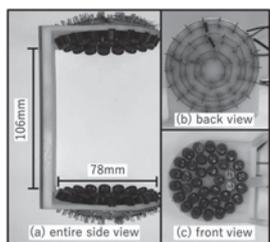


Fig. 1 Photographs of ultrasonic speaker units.

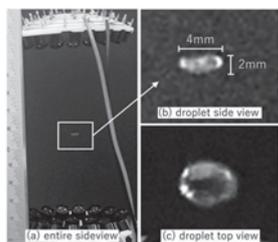


Fig. 2 Photographs of droplet capture.

## 2-7-14

### 2-7-14 矩形のステータとロータを放射状に配置した場合の特性

—近距離場音波浮揚による非接触ステップング搬送(16)—  
Characteristics when the rectangular stator and rotor are arranged radially, -Noncontact stepping transportation using near-field acoustic levitation(16)-

☆井上直幸, 孔徳卿, 梶原秀一, 青柳学 (室蘭工大・院)

- ◆本稿では円形に配置した振動子の屈曲振動を用いた非接触ステップング超音波モータ(NSC-USM)において、ステータを放射状にした際の動作について報告する。
- ◆扇形のステータとロータを用いた NSC-USM を Fig.1 に、矩形のステータとロータを用いた NSC-USM を Fig.2 にそれぞれ示す。矩形のモータは、ロータの停留位置の増加によるステップング分解能の向上が期待される。
- ◆二つの形状のモータにおいてロータに働く回転力の計算結果を Fig.3 と Fig.4 にそれぞれ示す。矩形のモータは扇形のモータと比べ、分解能が6倍に向上する。
- ◆実際に製作したステータを駆動したがロータの搬送は確認されなかった。

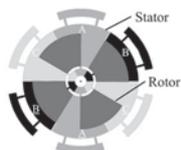


Fig. 1 Model of motor with 4-poles 60° rotor.

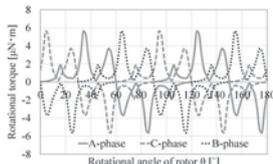


Fig. 3 Simulated rotational torque generated by phase switching in a fan-shaped motor.

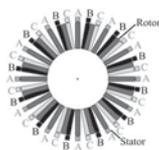


Fig. 2 Novel model of NCS-USM using radial array of rectangular vibrators.

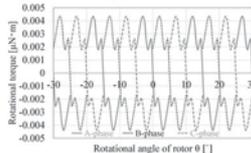


Fig. 4 Simulated rotational torque generated by phase switching in a rectangular motor.

## 2-7-13

### 2-7-13 たわみ振動を用いた振動板による粉体搬送

Powder transportation through a vibrating plate using flexural vibration

☆中岡夏海, 小松浩之, 小山大介, 松川真美(同志社大)

- ◆本研究では、たわみ振動板からなる粉体搬送路を用いて、その内部に発生する振動や音響定在波を利用した、粉体の均し搬送を目指している。本報告では、V字型振動板を用いた粉体搬送について検討した。
- ◆片栗粉と酸化マグネシウムを用いた粉体搬送において、搬送路長手方向に粉体表面の高低差を 600 μm 以内に収めることができた(Fig. 1)。しかし搬送路内に発生する定在波音場(Fig. 2)により、粉体が中央部に移動し、短手方向には 2.4mm もの高低差が生じた(Fig. 3)。
- ◆今後は、位相差駆動や振動板形状の最適化を行い、粉体搬送特性の改善を目指す。



Fig. 1: Powder transportation by ultrasound vibration.

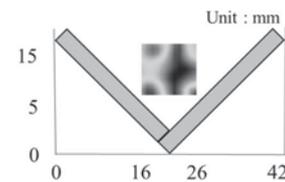


Fig. 2: Sound pressure distribution in the V-shaped plate.

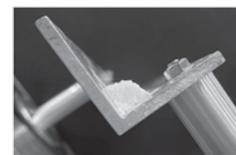


Fig. 3: Cross-sectional view of the transported powder.

## 2-7-15

### 2-7-15 振動面上から跳躍現象が生じる際の板状物体の過渡挙動

The transient behavior of a tabular object at that time the jumping phenomenon from a vibration surface

◎青野浩平, 青柳学(室蘭工大・院)

- ◆板状物体が跳躍現象により振動面上から垂直方向に跳躍する際の浮揚距離をレーザー変位計により測定した。
- ◆Fig.1 に使用した振動子の一例を示す。振動面の直径が 50mm であり、深さ 3mm の逆円錐形となっている。また、約 39.7kHz のとき同位相で中心から外側に向かって振動振幅が大きくなる。
- ◆Fig.2 に Fig.1 の振動子を用いた時間に対する浮揚距離 H および振動振幅 A を示す。ここで、浮揚物体は直径 50mm、厚み 0.1mm および 0.402g のポリエステルシートである。また、振動子に周波数 39,760Hz および 80, 120Vpp で印可したときを比較した。Fig.2 より H が約 1.5mm に達すると跳躍することが分かる。また、電圧振幅増加により跳躍するまでの時間が短縮されることが分かった。

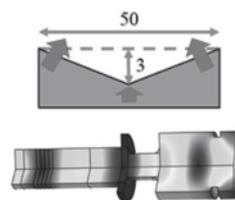


Fig. 1 Inverted conical recessed surface (39.7kHz, unit: mm).

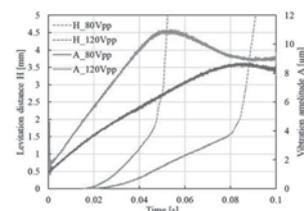
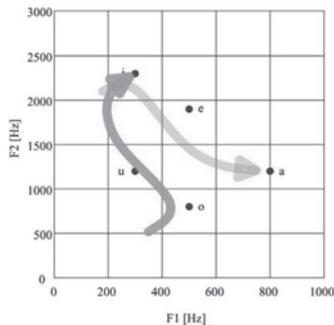


Fig. 2 Time vs. levitation distance H and vibration amplitude A (80, 120Vpp, 39,760Hz).

## 2-8-1

2-8-1 インタラクティブ音声合成システムにおける  
ガイドラインの動的表示に関する一検討A Study on the Dynamic Display of Guidelines  
for an Interactive Speech Synthesizer☆後藤 大樹(北海学園大), 青木 直史(北大),  
藍 圭介(北大), 元木 邦俊(北海学園大)

身体動作を駆使した発声代行は、ユーザーの操作の「くせ」によって一種の個性を表現できる可能性を秘めているが、まったくの初心者か何の手がかりもなく自由自在にフレーズを合成することは難しい。本研究では任意のフレーズから調音結合を考慮したガイドラインを自動生成し、動的に変化するガイドラインをグラフィカルに表示する手法について検討している。



「こんにちは」のガイドライン

Voice Padは「<https://bit.ly/2Wuuux3>」に公開しています。

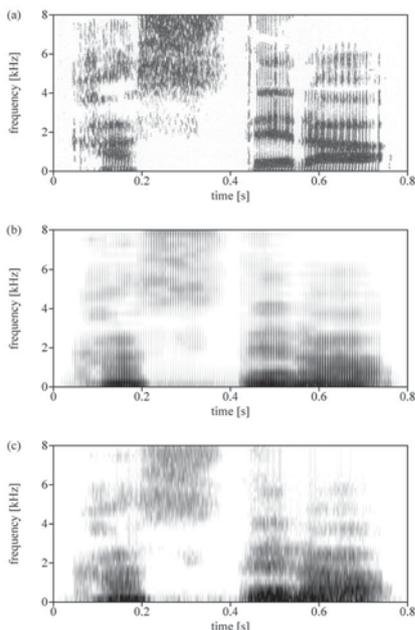
## 2-8-3

2-8-3 デジタルパターンプレイバックにおける  
有声/無声判定に関する一検討

A Study on the Voiced/Unvoiced Decision of Digital Pattern Playback

☆大竹 裕季, 青木 直史, 尾岡 剛成, 土橋 宜典(北大)

本研究では、デジタルパターンプレイバックによって復元した音声の聞き取りやすさの改善について検討している。



デジタルパターンプレイバックによって復元した音声

## 2-8-2

2-8-2 パーソナライズされた音声合成のための  
日本語音素バランス文作成に関する一検討A Study on the Creation of Japanese Phoneme Balanced Sentences  
for Personalized Speech Synthesis

☆高井 佑洗, 青木 直史, 土橋 宜典(北大)

ALS(筋委縮性側索硬化症)や咽喉摘出手術などによって声を失うことが不可避である場合、あらかじめ声を録音することによって、本人の声をういた音声合成を実現できるようになってきている。

こうした目的のための録音用原稿として最適な音素バランス文を生成することができるアプリケーションを作成することができれば、ユーザーのQoL向上にとって大きく貢献する可能性がある。本研究では、こうしたユーザーの要求に適した音素バランス文を作成するためのアプリケーションに必要な機能と、それを達成するための方法について述べる。

000: スプリットイメージや二重像の重ね合わせによりピント合わせを行う  
001: 寿命の到来を悟ったバットンは突如デッキカードを救い上げ  
002: マーラーは序曲を含むオペラ上演中の客席入場を禁じた  
003: ヒンドゥー教や大衆仏教を信奉してきた東南アジアの王朝では  
004: 結果として糖尿病などの成人病的症状で動物病院に通院するペットがいる  
005: 葡萄茎をほぐして株を分け一定間隔で植え付ける植芝という方法もある

図1 生成された音素バランス文の一部

A part of created Japanese phoneme balanced sentences

## 2-8-4

## 2-8-4 日本語音声の diversity とその展開

Diversity of Japanese spoken language and its perspectives

○白勢 彩子(東京学芸大)

## はじめに

日本語学の領域で論じられてきた、日本語の多様性の議論の経緯を振り返り、日本語音声の多様性を捉える観点を整理する。これらをふまえ、領域間の接続、研究以外の一般社会への還元・展開について論じる。

## 日本語の多様性—「位相」

同じ話題でも、相手や場面により選択する言語表現は異なる。日常使うことばはバリエーションが豊かであり、ことばの運用を考える上で、不可欠な話題である。特に日本語は、敬語や話しことばと書きことばのような、場面による差や地域による差が明確である。これらの相違を、日本語学では「位相」、また表現形式を「位相語」と呼んでいる。

## 日本語音声における位相・変異

音声言語の位相研究に目を向けると、地域差、つまり方言を扱うことが多いことに気が付く。音声言語の性別や年齢による差は、「話者の個性」のキーワードで、音響学領域で研究成果が蓄積されてきている。感情、パラ言語の音声表出も、相手・場面による差である。これについても、音響学領域での進歩が目覚ましい話題である。音声の多様性の研究の進展を考えると、領域間の融合が非常に重要であることが明確となり、音声によるコミュニケーションの研究の広がりが期待される。

### 2-8-5

#### 音声コミュニケーションのダイバーシティ —日本の方言音声について—

Diversity in Speech Communication: Phonetic Variation in Japanese Dialects

○木部暢子(国語研)

◆日本各地の方言音声の多様性について報告する。母音、子音、質問文の文末音調の地域差を取り上げ、多様性が生じた過程を考える。

◆母音の多様性 最も母音の種類が少ないのは、沖縄県と那国の3母音、最も多いのは愛知県名古屋市の8母音、その間の4~7の母音を持つ方言が各地に分布している。母音の数が少ない地域は母音間に統合が起きたもの、母音の数が多地域は連母音の融合により新たな母音が生じたものである。

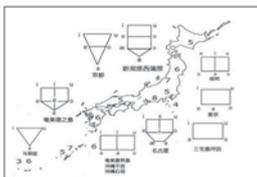


Fig.1: Number of vowels

◆子音の多様性 ハ行子音に対応する子音のバリエーションが琉球で豊富である。ハ行子音は、古く(奈良時代)は近畿で \*p と再建されている。その後、多くの地域で \*p > \*ɸ > h の変化が起きた。琉球には、この3段階の状態が分布している。

◆質問文の文末音調の多様性 質問文は一般に文末が上昇すると言われているが、各地方言には4タイプの文末音調が観察される。(A)WH 疑問文、Yes/No 疑問文いずれも上昇、(B)WH 疑問文、Yes/No 疑問文いずれも下降、(C)WH 疑問文は下降、Yes/No 疑問文は上昇、(D)WH 疑問文が文末に向かって漸次上昇。文末の音調は、文中に質問の意味を担う文法形態素が存在するか否かと大きく関係している。

### 2-8-7

#### 2-8-7 人と機械の文字起こし能力比較: 自動音声認識は人間を超えたか?

Humans versus machines: Who are better transcribers?  
—Observations from Japanese transcription tasks

○加藤宏明, 河井 恒, 水上悦雄 (NICT)

- ◆機械による自動音声認識 (ASR) 性能の現状を人の音声書き起こし能力との比較において示す。
- ◆ASR 評価用の日本語音声データセットを対象として、書き起こしの能力を、速さと正確さの両面で競わせた。
- ◆人の書き起こしの単語誤り率を1ストローク毎に逐次モニタするツールを作成して精密かつ公平な評価を行った。
- ◆速さでは、ASR は人の2倍あるいはそれ以上の性能であり、機械が人を大きく上回った。正確さでは、両者ともに平均単語誤り率が3%未達であったが、ASR が誤認識する少数例 (Fig.1) にも正答する人の能力がまだわずかに機械を上回る結果となった。

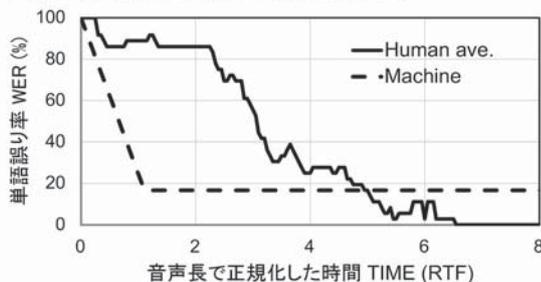


Fig.1: Word error rate (WER) with time in transcribing an utterance by humans and machines. Machines rapidly achieved a fairly good accuracy. Humans improved accuracy very slowly but finally showed the better performance.

### 2-8-6

#### 2-8-6 World Englishes と英語ペラペラの間で音 声学を教えるということ

Teaching phonetics at the frontline between World Englishes and native-like English speech

○北原真冬(上智大)

英語を母語としない私たちの音声コミュニケーションにおいて、英語という言語にどう対峙するかは、避けて通れない課題である。(a)初等・中等教育を受けてきた立場から学習対象(受験科目)としての英語、(b)音声を研究する立場から研究対象としての英語、そして(c)その研究内容を発表・公開する立場から道具・媒体としての英語、という少なくとも三つの英語像を、私たちのそれぞれが持っているであろう。さらに、筆者のように大学における英語教員という立場にあると、(d)教授対象としての英語という像が加わる。本稿ではWorld Englishes という概念について簡単に紹介した上で、特に(b)と(d)の英語像のあり方について、音声学・音韻論を専門とする立場から考察する。規範主義と記述主義という言葉に対する二つの見方について、少なくとも科学的に音声を分析する営みは規範主義とは相容れない、という前提に基づき、大学の英語学科に入学してくる学生に対して、「うまい英語」とは何かを自ら考えさせるように、音声学および英語を教えるという筆者の取り組みを紹介する。

### 2-8-8

#### 2-8-8 話者性に見る音声の多様性

Speaker characteristics and diversity in speech

○網野加苗(科警研)

音声に含まれる発話者に関する情報を話者の個性(話者性)と呼ぶ。話者の個性は、その人物を構成するさまざまな要因によって作られる。本発表では、話者の個性の種類と個性を変化・変動させる要因について、話者プロファイリングの観点から概観する。

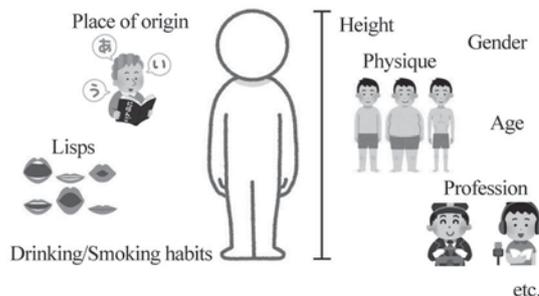


Fig.1. Speaker characteristics that appear in speech sounds



Fig.2. Factors affecting speaker characteristics

### 2-8-9

#### 2-8-9 人間の多様性を認め合う音声コミュニケーション

Audio communication technology for human diversity

○川森雅仁(慶大)

- ◆国連のSustainable Development Goals (SDGs)のモットーは「誰も取り残さない」である。本稿では、「誰も取り残さないために」人間の多様性(性別、年齢、障害の有無など)を許容するコミュニケーションに音声(技術)がどう関わるか、について述べる。
- ◆音声を聴けないため聴覚障害者が、電話相手の聴者の音声を映像(手話)あるいは文字に変換(モード変換という)してコミュニケーションを補助する仕組みが電話リレーサービスである。現在、このモード変換を完全に機械化できないため、主に人間が変換を行っている。しかし現在、自動音声認識、TTS、声質変換などの技術が研究されている。
- ◆このような電話リレーサービスは、これまで、わが国では提供されていなかったが、2020年に新たに法律が成立・施行され、公的インフラとしての電話リレーサービスが2021年から提供される。
- ◆可聴範囲を超え、音を振動であると広く考えると、聴覚障害者にとっても音自体が重要なコミュニケーション手段となりうる。その代表的なものが、聴覚障害者向けのハプティックによる音楽体験といえる。
- ◆音声による視覚障害者の歩行支援は以前から重要な適用領域である
- ◆スマートグラスやスマート耳栓のようなウェアラブル機器の発達の結果、視覚障害者のための超音波を使った位置同定システムが現実的になってきた。
- ◆人間の感覚にとって本質的な音をより多様に利用することで、音声コミュニケーションはSDGsにとっての重要な技術要素となると思われる。

### 2-9-1

#### 2-9-1 三線における胴の振動と音色の関係性の評価

Evaluation of the relationship between the vibration of sound box and the timbre of *Sanshin*

☆杉本理乃, 西宮康治朗, 長秀雄(青山学院大)

沖縄の伝統楽器である三線は胴に棹が差し込まれた構造をしており、棹の振動によって胴の振動が変化し、音色に影響を与えると考えられる。そこで棹の振動が胴を介して音色に及ぼす影響を解明することを目的として、棹の有無による三線の実験モード解析および有限要素法解析を行った。実験モード解析結果より、棹の振動によって木棹の振動が変化することが分かった。また、音の振幅は棹の有無における木棹の変形量によって変化した。有限要素法解析結果より、膜の振幅が大きくなったモードでは棹の振動によって木棹が大きく変形し、木棹と膜が連成振動を起こして膜が大きく振動した。以上より、木棹と膜のモード形状の関係が音色への影響に重要であることが分かった。

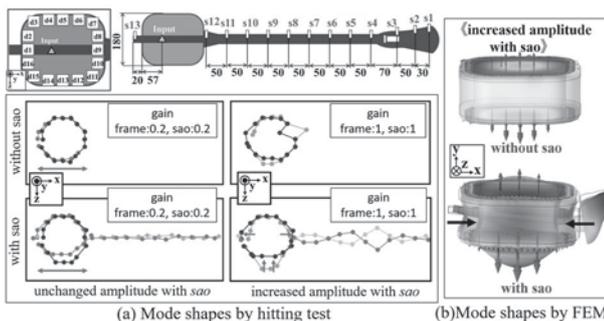


Fig.1 Results of impact test and FEM on *Sanshin* sound box with and without sao

### 2-8-10

#### 2-8-10 発声発語訓練のための視覚的フィードバックに向けた聴覚障害者による発話の音響分析

Acoustic analysis of the speech produced by deaf or hard of hearing people for visual feedback at speech training

○安啓一, △石井悠貴, ▽種子田尚人(筑波技術大)

- ◆聴覚障害者の中で、特に音声言語獲得する前に失聴し、補聴器での聴覚活用が困難なろう者は自身の声がフィードバックされにくいことから、発話・発音が不明瞭になるケースがある。
- ◆現行の発音訓練に音声分析ソフト(モニター画面)による視覚的フィードバックを加えることで、よりよい発音訓練につながるのではと考え手始めとして朗読音声と自由会話(自発語)音声分析ソフトを用いて各母音・子音の分析を行い、これらの特徴量が発声発語の訓練の補助になるかどうか予備的な検討を行った。
- ◆聴覚障害者7名(19-23歳, 平均22.0歳)と健聴者6名(19-23歳, 平均21.5歳)の計13名が実験に参加した。収録した音声は朗読音声(北風と太陽)と自由会話である。
- ◆ピッチについて個人内での比較を行ったところ、朗読音声よりも自由会話でピッチが低くなった(平均12.9Hz 低下・有意差あり)。聴覚障害者のピッチの変動幅について、男女いずれも健聴者と比べ減少を認めた(女性は平均35Hz減少, 男性は平均10.3Hz減少)。
- ◆フォルマント分析の結果、5母音の分布に群間差は認められなかった。
- ◆ピッチ及びフォルマントをリアルタイムにフィードバックができるソフトウェアを発声発語訓練の現場に持ち込み、発音訓練の補助になるかどうか検証を行ったところ、母音と子音の区間の判別や摩擦音の訓練で効果がある可能性が示唆された。

### 2-9-2

#### 2-9-2 音響特徴量を用いたヴァイオリンの演奏表現の違いによる音色の解析

Analysis of the timbre in performance expression of a violin by acoustic features.

☆石垣優弥(明星大), 横山真男(明星大)

- ◆本研究では、音響特徴量の音色表現における変化の解析を行い、音響特徴量が音楽的な表現とどう結びつくのか解析した。
- ◆使用する楽曲はヴァイオリンの音色表現(Powerful, Rich など)に似合いそうなクラシック音楽の楽曲を選定した。
- ◆実験では音色表現あり/なしで引いてもらい、各音響特徴量を抽出した。
- ◆各音響特徴量の値から音色表現と音響特徴量の考察、また音響特徴量の時間変化に現れる演奏法における違いの考察を行った。
- ◆以下の図は音色表現 Powerful のスペクトル重心の表情あり/なしにおける時間変化の様子を表したものである。

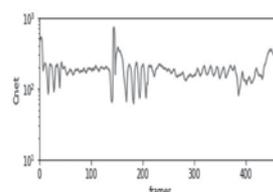


Fig.1: The time variation of Centroid with expression

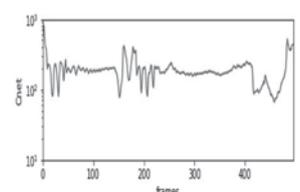


Fig.2: The time variation of Centroid without expression

### 2-9-3

#### 2-9-3 ヴァイオリン演奏における弓の踊りの数値解析

Numerical analysis of tremble of the bow in violin performance

○鮫島俊哉(九大・芸工)

- ◆擦弦振動の等価回路表現に、有限要素法による弓の振動場の解析モデルを導入し、擦弦振動場と弓の振動場が連成された解析手法を構築する。Fig. 1は、解析対象とする振動システムを表している。
- ◆Fig. 2のような解析事例により、ヴァイオリン学習者が抱えている課題：「弓が踊る(震える)」現象の発生機構を解明する。

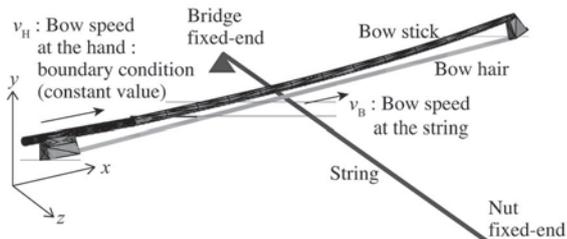


Fig. 1 : Geometrical arrangement of the bow and string.

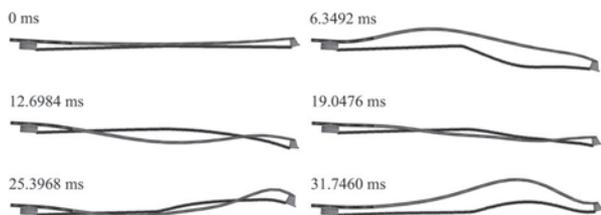


Fig. 2 : Snapshots of the time evolution of the shape of the violin bow.

### 2-9-5

#### 2-9-5 スネアドラム演奏における演奏正確性と手動作の特徴との対応関係

Correlation between performance accuracy in snare drum performances and characteristics of hand movement

☆小出英範(東京情報大院), 西村明(東京情報大)

- ◆東京情報大学の学生 11 人 (ドラムス経験者 5 人、未経験者 6 人) を対象にスネアドラムの連続打拍の演奏実験を実施。
- ◆演奏時は先行研究で分析された演奏テンポの差異の他に、両手の演奏と片手の演奏での仕事量が異なるデータも収集。
- ◆演奏音から演奏リズムの正確性を評価する特徴量を抽出。
- ◆動作データから手の振りの波形の再現性を評価する特徴量を抽出。
- ◆波形の再現性の分析は、1 周期でのサンプル数や振り上げ高さに対して補正がある場合とない場合でそれぞれ分析。
- ◆上記の特徴量の相関分析を行うことで、正確なリズムで演奏している演奏者の演奏動作の特徴を検討。
- ◆1 周期のサンプル数が無補正の場合では手の振りの波形が同じでも、手の振りの再現性と演奏リズムの正確性に相関関係が表れる。
- ◆手の振りの波形が一定になるように演奏すると一定のリズムで正確に演奏できることが示唆された。
- ◆手の振り上げ高さや振り下ろしの深さもリズムの正確性に影響する可能性があり、今後分析する必要がある。

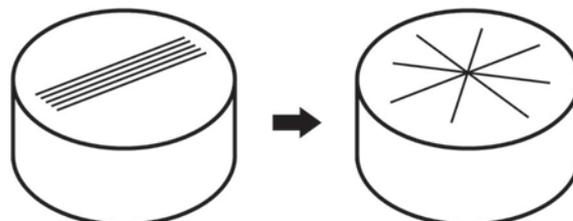
### 2-9-4

#### 2-9-4 スネアドラムのスナッピーの材質や形状が音色に与える影響の数値解析

Numerical analysis of the effect of changing material or shape of snare of the snare drum.

☆関口周, 鮫島俊哉 (九大・芸工)

楽器の形状や部品構成というものは、長年の経験によって完成されてきた面があるゆえ、それ自体に疑問をいだく人は少ないように思われる。それは本当に最良のかたちなのだろうか？今までに無い音色を作り出せるような楽器というものは、今後はもう作れないのだろうか？作れないのならばそれはなぜか？作れるのならそれはどのようなかたちをしているのだろうか？



### 2-9-6

#### 2-9-6 モーションキャプチャを用いた鍵盤演奏動作の教育用可視化システムの検討\*

\* Examination of educational visualization system for keyboard performance using motion capture

○三戸勇気(日大・芸術), 幸田優志(日大・理工), 川上央, 駒澤大介(日大・芸術), 三浦雅展(国立音大), 山口達也, 篠田之孝(日大・理工)

現在まで、モーションキャプチャシステムを用いて演奏動作の計測ならびに分析を行ってきた。演奏家の育成にとって、その計測したデータや分析した結果を演奏者にわかりやすく示すことが、より有効的なデータの活用法だと考えた。そこで、今回我々は演奏動作データをわかりやすく提示することを目的として、鍵盤楽器演奏における教育用可視化システムを構築した。

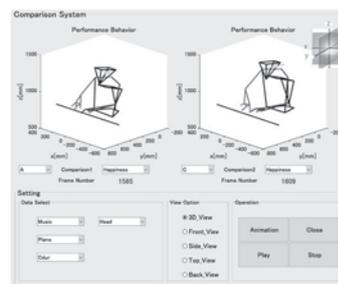


Fig.1 Educational visualization system for keyboard performance

## 2-10-1

## 2-10-1 【招待講演】音バリアフリーの観点から視覚障害・聴覚障害者のオンライン会議ツールの利用状況を考察する

A discussion of the usage of online meeting tools for the visually and hearing impaired persons from the standpoint of sound barrier-free

○坂尻正次(筑波技術大学)

- ◆視覚障害・聴覚障害者のための国立大学である筑波技術大学において授業等で用いたオンライン会議ツールの利用状況について考察した。
- ◆特に音声の利用が難しい聴覚障害の方に懸念があったが、双方の学生とも概ねオンライン会議ツールを有効に利用できた。
- ◆視覚障害・聴覚障害で音声を利用している場合、オンラインでのマイクを通じた音声になることで聞き取りやすくなったとの声があった。
- ◆図1は、聴覚障害学生向けにOBSにより手話動画と字幕、さらには手書きも可能な状態でZoomを用いて配信している様子



Fig. 1A distribution screen of Zoom using OBS for the hearing impaired students

## 2-10-3

## 2-10-3 (招待講演)コロナ禍におけるオンライン授業での障害学生支援の現状～聴覚障害学生への支援の現状～

Supporting in Online Classes for Students with Hearing Impairments in the Time of COVID-19

○三好茂樹

- ◆本年初頭からのコロナ禍において、オンライン授業が教育機関で開始されている。このようなオンライン授業において、聴覚に障がいのある学生の修学環境にも変化が起きている。今回の講演において、これらの修学環境改善のための相談・支援の事例等について触れたい。

## 2-10-2

## 2-10-2 (招待講演)コロナ禍におけるオンライン授業での障害学生支援の現状～視覚障害学生への支援の現状～

Supporting in Online Classes for Students with Visual Impairments in the Time of COVID-19

○宮城愛美(筑波技術大)

- ◆近年、情報アクセス技術の向上と、法制度の整備を始めとする社会の動きにより、視覚障害学生の修学環境に改善がみられる。コロナ禍によって、これらの環境にどのような変化が生じたのか。本講演では、まず、視覚障害学生の修学と全国の大学等で行われている支援について触れ、その後、コロナ禍における視覚障害学生に対する教育・支援の状況と課題について述べる。

## 2-10-4

## 2-10-4 重度身体障害者のオンライン会議ツール活用の工夫からみる音声支援

Audio support approach seen from the practical use of online meeting tools by people with severe physical disabilities

◎藪 謙一郎(東大)

- ◆外出が困難であった重度障害者は、近年では映像音声通信技術やロボットにより仮想的な「外出」が可能になりつつある時代背景にあった。
- ◆そのような中、多くの人々が外出制限を迫られ、オンライン会議などのツールが急速に誰もが知るツールとなり、障害の有無に関係なく、各地でオンライン会議ツールが使われるようになった。
- ◆本稿では、重度身体障害者自身がオンライン会議ツールを活用する際に、工夫された次の3つの例について紹介する。
  - 構音障害に対して、音声と文字表示を併用させる使用方法
  - 人工呼吸器から発せられる雑音に対処した方法
  - オンライン会議ツールを使用した音楽活動

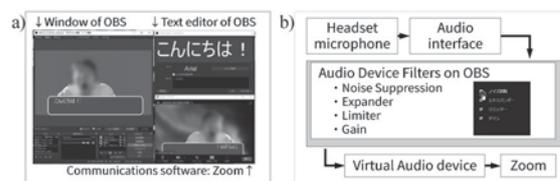


Fig. 1: Examples of text display for articulation disorders (a), and Sound filtering system for reducing ventilator noise (b).

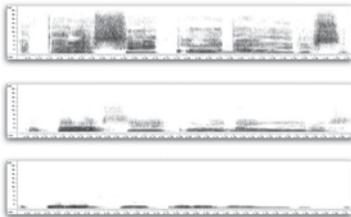
## 2-10-5

2-10-5

講義配信や飛沫対策用防護具による音質劣化が  
即時聴解に与える影響Influence of speech quality degradation by remote lecturing  
and wearing masks on instantaneous intelligibility

峯松信明, 西村多寿子, 箱田峻 (東大・工)

- 2020年冒頭よりCOVID-19が流行。パンデミックに。
- 三密を避けるために生活習慣が一変。
  - ・社会的距離 (ソーシャルディスタンス) の徹底
  - ・マスク, フェースシールドなど飛沫対策用の防護具の装着など
- 教育・学術活動における大きな変化
  - ・授業・講義はほぼ全てがオンライン化
  - ・学会活動もオンライン化。全国大会, 研究会, 国際会議も同様
  - ・ハイブリット授業になると, マスクをした上で講義配信
  - ・通信条件やマスク装着によって, 音質は劣化することに!!
- 通信条件やマスク装着に起因する即時聴解の崩れを分析
  - ・4種類のマイク x 3種類のマスク条件 with Zoom で音声収録
  - ・聴解の崩れをシャドーイング法によって定量化
  - ・日本人は大丈夫でも, 留学生はどうなる!?



## 2-11-1

## 2-11-1 細野晴臣音楽作品の3D化の試み

An attempt to convert Haruomi HOSONO's music into 3D audio

岩田渉 (一般社団法人オブジェ・アー)

- ◆ 2019年10月に1か月間行われた細野晴臣デビュー50周年記念展「細野観光 1969-2019」において細野氏が50年にわたって製作してきた音楽作品を3D化して展示した。長時間聴いても飽きない居心地のよい音楽空間として評価された。その3D音響デザイン手法のノウハウについて紹介する。

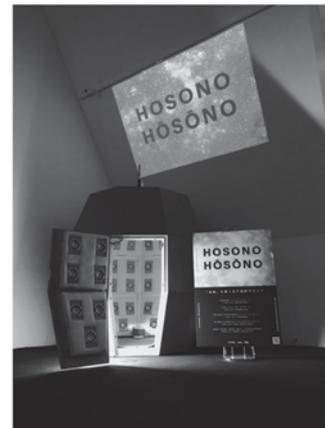


Fig.1: Installation view

## 2-11-2

2-11-2 (招待講演)物語世界における  
主観的な聴取点デザインの諸課題

Some Challenges toward Designing for "Point of Audition" in Narrative World

○伊藤彰教(東京工科大・メディア)

- ◆ 「映像コンテンツの視聴中、この音をわたしは物語世界のどこから聴いていると考えればよいのでしょうか?どこに位置付けるのが適切ですか?」
- ◆ 「このカメラワークは登場人物の「〇〇目線の映像だな」と分かりますが、音についてはどうでしょうか。「〇〇目線」にあたる聴覚の用語ってありますか?」
- ◆ 「アンビソニックマイクで収録した音声データをゲームに組み込んでみました。プレイヤーキャラクターがゲームの世界の中を動くのですがアンビソニックマイク収録の音はどのように位置させるべきでしょうか。このソフトウェアの設定はどう操作すればよいですか?」
- ◆ 「バイノーラルマイクで『わたしに接している登場人物が話しているような音』は録ることができますか?」
- ◆ 「カメラ位置に置いたマイクだと音が小さいし遠いです。なぜですか?カメラ位置にマイクを置いたらダメな理由は何かですか?」
- ◆ 「アンビソニック収録したライブをVRで楽しんだら臨場感がすごく良さそうと思って作ってみたのですが、首を振ったり自分の位置を変えたら音楽が回って気持ち悪いです。音楽だけ常にステレオ固定したらダメですか?」
- ◆ 本講演は、日々このような質問を学生から受けて答えに窮する教員の苦闘について講演します

## 2-11-3

2-11-3 3D オーディオによって彩られるゲームの世界  
— 実際の実装と最近の話題を通して —

The world of games composed of "3D-Audio": Consideration with actual sound designing and topics of audio processing recent-day

○小塩広和(フリーランス), △田中孝(MyDearest)

- ◆ 本講演ではゲームにおける3Dオーディオのデザインについて、筆者が最近制作に参加したVR(ヴァーチャルリアリティ)アドベンチャーゲームALTDEUS: Beyond Chronos(以下ALTDEUS)について実例を交えて紹介する。また、最新のゲーム機に搭載されたオーディオ技術についても紹介する。
- ◆ ゲームではサウンド処理を出来るだけ軽量化する必要があり、物理法則よりも機能性を優先したサウンドデザインも必要である。
- ◆ ALTDEUSでは2D(距離、方向のみを反映)の配置を基本として、リアルタイムに残響を付加し空間を演出した。また、ボイスなど重要な音については距離減衰をカットした。BGMにボイスや効果音が消されないように最小限のダッキング処理を行った。
- ◆ 次世代のゲーム機では処理能力が大きく向上し、頭部伝達関数(HRTF)をリアルタイムで畳み込むことが可能になったため、ヘッドフォン/イヤフォンでリアルタイムに立体音響の再生が出来るようになった。
- ◆ このことでゲームのサウンドはより現実に近い表現が可能となるだけでなく、低コストで立体音響のコンテンツを制作することが可能になると思われ、今後多くの3Dオーディオを用いたゲームコンテンツが創出されるだろう。

## 2-11-4

### 2-11-4 音楽制作における3Dオーディオの可能性

The Potential of 3D Audio in Music Production

○亀川徹(東京藝大)

- ◆音楽は古くから演奏される空間と密接な関わりを持っており、立体音響としての効果を考えられてつくられた作品も数多くある。
- ◆22.2マルチチャンネル音響などの3Dオーディオは、空間を生かした音楽の表現をおこなうことが期待できる。
- ◆ここ数年HOA(高次アンビソニックス)がPCの制作環境でも可能となり、DAW(Digital Audio Workstation)のプラグインとして使用することで、音楽制作にも広く用いられるようになった。
- ◆本学で取り組んでいるチャンネルベースやHOAを用いた手法を例にあげながら、音楽制作における3Dオーディオの可能性について考察する。

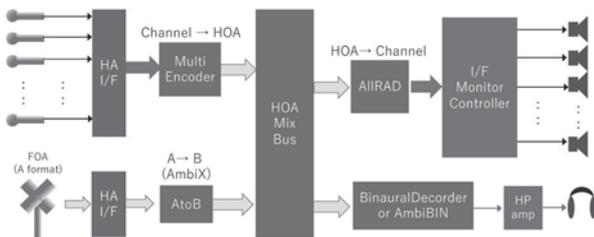


Fig.1: HOA(高次アンビソニックス)信号をマスターとしてミキシングする場合の信号の流れ

## 2-11-6

### 2-11-6 360度パノラマ映像と融合した音響コンテンツ制作

Sound content production with 360-degree panoramic video

☆柏崎 紘, 尾本章(九大芸工)

- ◆24チャンネル音場再生システムに、円筒スクリーンによる360度パノラマ映像を付加し、没入型視聴覚ディスプレイを構築した。
- ◆映像は全天球カメラを使って収録し、円筒スクリーンに合わせて歪み補正を施し、4台のプロジェクタにより投影する。
- ◆視聴覚融合に関する実験やデモンストレーションに使用するため、日頃よりコンテンツ制作を行っている。
- ◆デモでは収録点が移動するコンテンツが人気。移動音源や反響が変化するシーンは、音場再生システムの特徴が伝わりやすい。これまで「音だけ」では状況が伝わりづらかったシーンも、映像が加わることで楽しんで視聴してもらえるようになった。

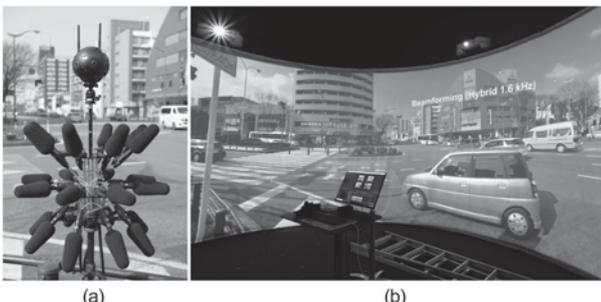


Fig.1: (a) Hedgehog microphone and Spherical camera, and (b) 360-degree panoramic projection onto cylindrical screen.

## 2-11-5

### 2-11-5 3Dオーディオにおける収録法と表現のデザイン

Recording methods and design of expression in 3D Audio

○入交英雄(WOWOW)

- ◆制作技術者の立場から3Dオーディオの表現について考察する。
- ◆クラシック録音に立ち塞がる「明瞭度」と「響き」の両立
  - 多チャンネル再生で可能な方向性マスクングの解除を利用し、2ch再生で解決困難な同時マスクングや経時マスクングの影響を軽減
  - 多チャンネル再生では、残響成分など2ch制作では聴こえ難かった音が聴こえるようになり、明瞭性と豊かな響きの両立が実現
- ◆3Dオーディオの重要な特徴は包み込まれ感と拡がり感
  - 必要条件であって十分条件でない。
  - 迫力などの音色要素が重要
- ◆収録の要素を、オブジェクト臨場感とフィールド臨場感に分ける
- ◆オブジェクト臨場感の向上=楽器そのものの表現
  - 音源の収録をモノマイクからマイクアレイ(点から面)へ拡張する
  - 楽器自身の縦と横への拡がり表現することで臨場感が高まる
- ◆フィールド臨場感の向上=包み込まれ感と拡がり感の表現
  - 間接音を3次元のマイクアレイ(キューブアレイなど)で収録。
  - 聴取者を取り囲むように配置したスピーカで再生
  - つながり良い「包み込まれ感」を表現することで臨場感が高まる
- ◆本例は表現法のほんの一例
  - 3Dオーディオ収録の普及→収録法の一般化が必要
  - 立体音響のための音響心理研究にアカデミズムによる解析を期待

## 2-11-7

### 2-11-7 3次元音場再生システムにおける再生方式の検討と映像付加による主観的評価への影響

A Study on the Reproduction Method of 3-D Sound Field System and the Effect of Adding Video Images Subjective Evaluation

☆渡壁航平, 尾本章(九大芸工)

- ◆3次元音場再生システムの低音過剰収録を補正する音場再生方式の適用及び映像付加が、聴取印象にどのように影響するかを調査した。
- ◆実験は、6つのコンテンツについて行われた。
- ◆実験の結果、音場再生方式はどのコンテンツでも一様に評価の高い再生方式はなく、呈示するコンテンツによって好まれる再生方式が異なることがわかった。また、映像を付加した時聴取体験がより好まれた。

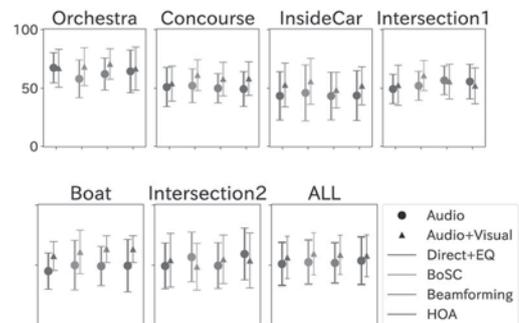


Fig.1: Result on the attribute 「好き」

### 2-11-8

#### 2-11-8 Multiple Vertical Panning を用いた 立体音響システムにおける 奥行き表現の比較検討

Comparative study of the depth expression  
in 3D audio system using multiple vertical panning

○木村敏幸(東北学院大)

- ◆我々はこれまでに大画面ディスプレイに適合した立体音響システム (Multiple Vertical Panning, 以降「MVP方式」と呼ぶ)を提案し、実用化の可能性も示してきた。さらに、MVP方式による音の奥行き表現も検討してきたが、奥行き表現に関するこれまでの研究ではステレオホニックのような従来方式とは比較してこなかった。
- ◆本報告では、MVP方式による奥行き表現の可能性を検討するために、ステレオホニックの音も再生できる視聴覚提示システムを制作し、臨場感を評価した。その結果、Fig. 1に示すように、映像が中心にあり視聴位置が中心位置以外の場合、MVP方式はステレオホニックよりも臨場感が向上することが分かった。また、点音源の物理的な音量変化を模擬すればさらに臨場感が向上することも分かった。

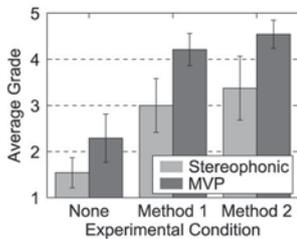


Fig.1: Results of the evaluation experiment.  
(Error bar: 95% confidential interval)

### 2-1P-2

#### 2-1P-2 雑音除去を用いたノンレファレンス 単語音声了解度推定法の検討

On non-reference speech intelligibility estimation using noise reduction.

☆上嶋雄大, 近藤和弘(山形大院・理工学研究科)

- ◆レファレンス信号に雑音除去によって推定した原音声を用いるフルレファレンス了解度推定手法を考案。
- ◆原音声の推定にDeep Learning 界隈で有名なGANを使用。
- ◆周波数重み付きセグメンタルSNRから了解度を推定するDNN(Deep Neural Network)を学習。
- ◆交差検定での推定精度を雑音種ごとのRMSEとピアソンの積率相関係数で評価を行った。
- ◆テストデータでRMSEは0.1455、相関係数は0.9032。
- ◆使用した主観了解度の分布の偏りがRMSEと相関係数に大きく影響。
- ◆ターゲットとなる主観了解度が低い場合、精度が低くなる可能性がある。

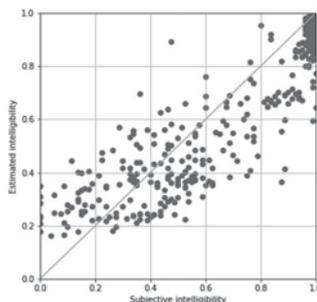


Fig.1: Distribution of subjective vs. estimated Intelligibility with test data.

### 2-1P-1

#### 2-1P-1 屋外での長距離伝搬条件で計測した HRIRを畳み込んだ拡声音声の了解度評価

Speech intelligibility evaluation  
using outdoor long-distance transmission HRIRs.

◎小林洋介(室蘭工大), 野口啓太(室蘭工大/TOA), 栗栖清浩(TOA)

- ◆屋外拡声システムの主観評価のために屋外の拡声フィールドに多数の被験者を集めることは現実的ではない。
- ◆この問題を解決するために実際の屋外拡声器から650から1200m離れた実際の聴取を想定した距離を評価条件に設定してダミーヘッドを設置し、その両耳で拡声器からのインパルス応答を計測した。
- ◆また、収集したインパルス応答を親密度別単語了解度試験用音声データセット2007(FW07)に畳み込み、単語了解度的主観評価を行なった。
- ◆10名の被験者に対し、録音地点4箇所、スピーカ種3種の12条件についてインパルス応答と高親密度(親密度7.0-5.5)と低親密度(親密度2.5-1.0)を評価した。
- ◆その結果、分散は大きいものの、被験者平均値でスピーカ種の差は議論可能であることが示された。

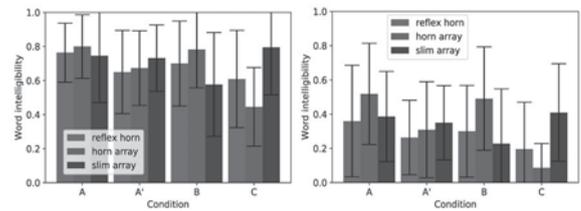


Fig.1 Word intelligibility by conditions

### 2-1P-3

#### 2-1P-3 オンライン会議ツールの 簡易脳波計測による評価の試み

An investigation of the on-line meeting tool evaluation  
with EEG measurement

○福島学, 沖田和久, 温水啓介, 山下涼介, 大里一矢(NBU), 大霜康晃, 加藤弘士郎, 石川直生, 上原正志(CIT), 河納隼一((株)アーネット), 近藤善隆((株)ジェイテック), 宇治川大吾(東京鋼鐵工業(株)), 窪田泰也(A-Lab.), 柳川博文(arsl)

- ◆オンライン会議ツールのパケット伝送遅延で生じるIRの変化に着目
- ◆対象ツール: Skype, Google meet, YouTube Live, ZOOM
- ◆短時間計測IR群(Fig.1)と音声から刺激音を作成
- ◆簡易脳波計により脳波(EEG)を計測
- ◆IRの時間構造変化に応じているような脳活動変化が生じることが確認できた

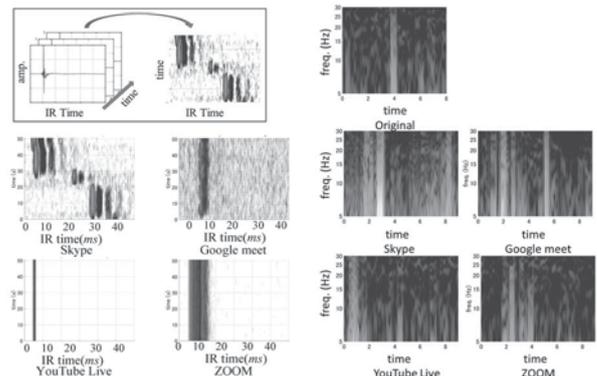


Fig.1 The changes of short-term measured Impulse Responses

Fig.2 Measured EEG with IRs convolved with voice

### 2-1P-4

#### 2-1P-4 残響劣化音に対する推定原音声を用いたノンレファレンス音声了解度推定

Non-Reference Speech Intelligibility Estimation Using DNN For Reverberation Degraded Sound.  
-Taking Into Account The Effect Of Individuality Of Speech-

☆中澤和司, 近藤和弘(山形大)

- ◆残響の影響を考慮するために「ここには〇〇と書いてある」といった文章に単語を埋め込んだテスト用音声を録音。
- ◆残響劣化させたテスト用音声に対するDRT試験を7名に実施し了解度を測定。
- ◆Bi-LSTMを用いて残響劣化音から原音声を推定。
- ◆推定した原音声と残響劣化音を用いてフルレファレンスを模擬した了解度推定のためのDNNを学習。
- ◆学習データはspeaker1(男性)の残響劣化音
- ◆未知話者 speaker2(男性)に対しても推定可能

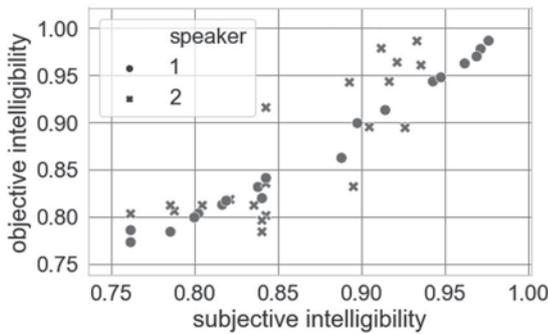


Fig. 1 Distribution of subjective vs. estimated speech intelligibility

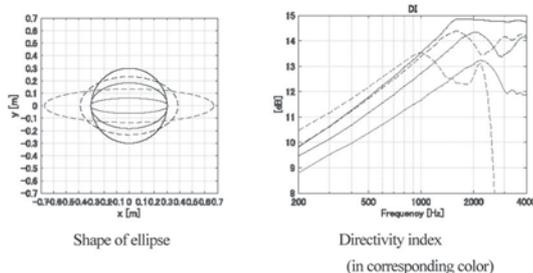
### 2-1P-6

#### 2-1P-6 剛楕円スピーカアレイを用いたビームフォーミングについての検討

Beamforming using Rigid Elliptical Loudspeaker Array

☆任逸, 羽田陽一(電通大)

- ◆アレイ信号処理技術のひとつとして、ビームフォーミング技術は通信や音響などの分野で広く研究されている。スピーカアレイの研究では、円形や球面アレイを用いることが多いが、筆者らは焦点音源生成技術において剛楕円スピーカアレイを導入する検討を行ってきた。
- ◆本研究ではビームフォーミング技術に注目し、最小分散無歪ビームフォーマを上述の剛楕円スピーカアレイに適用した際の、楕円の軸長や扁平率による再現性能への影響について検討する。
- ◆最小分散無歪ビームフォーマを用いて計算機シミュレーションを行った結果、片方の軸長が一定の場合、もう片方の軸長が長いほどビームの性能が上がることや、円周の長さが一定の場合、扁平率が低いほどビームの性能が上がる事が分かった。また、面積が一定の場合、扁平率の高い楕円アレイは狭い周波数帯域において鋭いビームを生成できることがわかった。



### 2-1P-5

#### 2-1P-5 SDMを用いた移動音源によって形成される音場の再現

Reproduction of the sound field generated by a moving sound source using SDM.

◎佐々木陽, 松井健太郎, 中山靖茂(NHK 技研)

- ◆任意の軌道で音源が移動する音場の角度スペクトル表現を導出した。
- ◆Firthらの方法 [1]で生じるエイリアシングの影響について、窓関数を取り入れることで回避した。
- ◆SDM (Spectral Division Method)を用いて、移動する音源によって形成される音場が再現できることを、数値シミュレーションにより確認した。

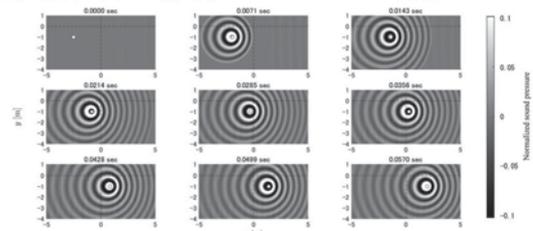


Fig.1: Desired sound field.

○: Sound source

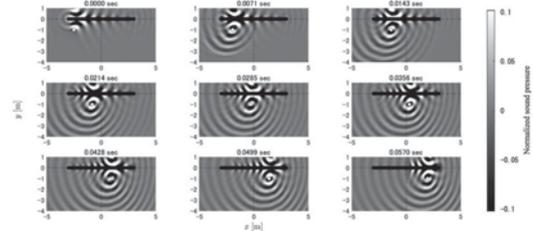


Fig.2: Reproduced sound field.

○: Virtual sound source, ●: Secondary sound source

[1] G. Firth and P. Ffiesl, JAES, 63(1), pp. 46-53, 2015

### 2-1P-7

#### 2-1P-7 円筒型スピーカアレイの内部における移動音像再生について

Reproduction of Moving Sound Image inside a Cylindrical Loudspeaker Array

☆久保健輔, 羽田陽一(電通大)

- ◆8素子の円形スピーカアレイを3段階積み上げた円筒型スピーカアレイの内部にて、移動音像の再生を試みた。球形を前提とするVBAPとHOA、及び最接スピーカ法(NS)とサイン則(SL)の4手法を用いて、音像の移動方向と滑らかさを評価した。
- ◆水平方向に音像を移動させたとき、全手法で移動方向の正答率は高く手法による差異は概ねなかった。滑らかさはHOAが良好であった。
- ◆上方に音像を移動させたとき、全手法で移動方向の正答率は高く手法による差異は概ねなかった。
- ◆螺旋状に音像を移動させたとき、HOAでの移動方向の正答率が著しく低い結果となった。滑らかさはVBAPが良好であった。
- ◆円筒型スピーカアレイにおいて、2次元ではHOA、3次元ではVBAPによる再生手法が有効であることが確認できた。

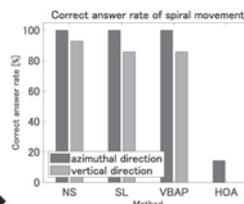


Fig. 1 Correct answer rate of spiral movement

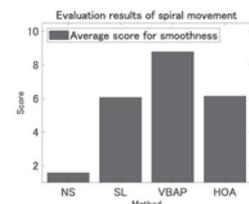


Fig. 2 Evaluation results of spiral movement

### 2-1P-8

#### 2-1P-8 平面マイクロホンアレイを用いた接話収録

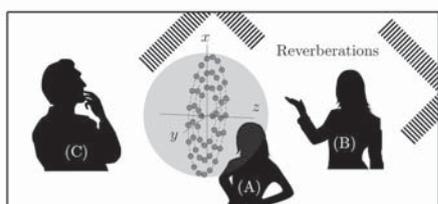
Close-talking recording using a planar microphone array

○岡本拓磨 (NICT)

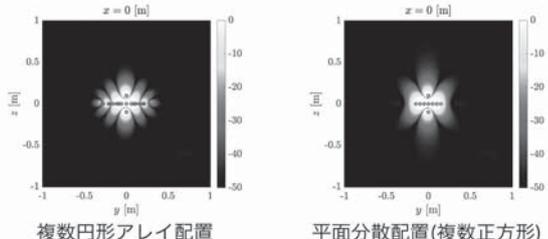
Q: 「平面上に配置したマイクロホンアレイで近くの音のみを収録したいのですが?」

オカモト: 「それならこれ!!」

- (a): z軸上に配置した鉛直マイクロホン対の収録音圧を
- (b): 複数の円形アレイ/平面分散アレイで推定
- (c): 差分音圧による3次元接話収録を実現



平面アレイを鉛直に回転させた使用イメージ



### 2-1P-10

#### 2-1P-10 受聴者に対する発話者の顔面の向きによる音声の周波数特性の変化について

Preliminary study on the radiation properties of uttered speech signals

☆清水一稀(秋田県大院)

安倍幸治, 西口正之, 高根昭一, 渡邊貴治(秋田県大)

- ◆人の音声を仮想的に再現するには頭部伝達関数(HRTF)以外に放射特性の影響を考慮する必要があると考えられる。本報告では、短い挨拶分や音韻ごとに放射特性の違いを確認できるか検討を行う。
- ◆短い挨拶文及び長母音をマイクロホンアレイで録音し、1/3 オクターブバンドごとに方向別にエネルギーを求め、それぞれレーダーチャートで作成した。
- ◆低い周波数では音韻ごとの放射特性の違いはあまり見られないが、周波数が高くなると発話者の横方向は弱く、正面方向に対しては強い指向性をもつなど、音韻ごとの放射特性の違いが見られるようになる。
- ◆複数の音韻が混ざると、放射特性は比較的同じとなる。
- ◆発話内容によって音韻ごとに放射特性が異なることを確認できた。

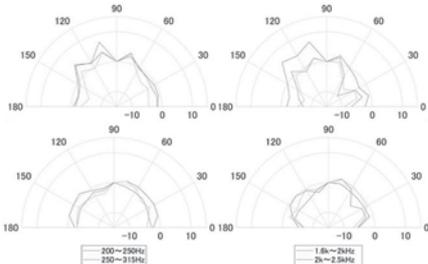


Fig.1: Radiation properties of the "Hello" (top), "A" (bottom)

### 2-1P-9

#### 2-1P-9 分散マイクロホンアレイのアレイ配置を考慮した仮想球モデル型広領域音場収録法の検討

Sound field recording based on virtual sphere model considering the arrangement of distributed microphone arrays

☆片田晃輔, 坂本修一(東北大通研/院情科研)

- ◆多くの場合、HOA 収録はマイクロホンアレイを聴取点に置いて行う。
- ◆しかし、実際には聴取点に機器が置けない場合が想定される。
- ◆我々は聴取点に機器を置けない場合を想定し、仮想球モデルを規範として、空間に複数配置したすべてのマイクロホンアレイから、仮想球上の音圧を最小二乗によって求める方法を提案した
- ◆本報告ではこの方法を改善し、空間に複数配置したマイクロホンアレイに対し、その配置と仮想球上のサンプリング点配置を考慮して求めた仮想球上の音圧から内部音場を表現する方法を提案する。
- ◆結果は、仮想球上の音圧の音圧を HOA エンコードし、音場を描画することで評価を行った

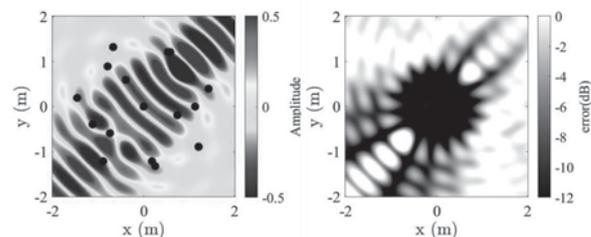


Fig.1: reproduced field (left) and error field (right) (black circles in the left figure represent the position and size of the microphone array)

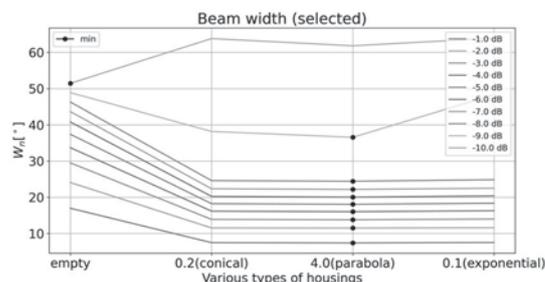
### 2-1P-11

#### 2-1P-11 FDTD 法によるマイクロホンアレイ筐体の影響を考慮した伝搬シミュレーション

FDTD Simulation for beamforming using various types of microphone array housings

○野口 賢一, 加古 達也, 中川 朗(NTT)

- ◆マイク周囲にディンプル型バツフルを形成するマイクアレイの伝達特性を、FDTD シミュレーションにより計算し、ビームフォーミング性能を改善するマイクアレイ筐体の検討を行っている。
- ◆ディンプル型バツフルとして、コニカル、パラボラ、エクスポネンシャルホーンを想定し、各ホーン形状に対して、広がりを制御する係数を振り、様々な筐体におけるマイクアレイのビームフォーミングの指向性を評価した。
- ◆指向性の鋭さを示すビーム幅 (Fig. 1) および、周囲360度任意方向に形成する指向性のビーム形状が等しいかを示すビーム平均偏りを評価し、ビームフォーミング性能を最も改善する筐体形状、パラメーターを求めた。



## 2-1P-12

### 2-1P-12 空間導関数標準化における微分係数の差分近似の影響

Influence of difference approximation on spatial derivative sampling

☆井上和弥, 島内末廣, 青木茂明(金沢工大)

- ◆ 離散点の音圧だけでなくその高次の空間的な微分係数を取得するマイクロホンアレイを構成すれば、空間標準化定理により規定されるマイクロホンの実効的な素子間隔は、微分係数の次数  $R$  に応じて、 $R+1$  倍に拡張可能である。
- ◆ 空間導関数標準化は、近接したサブアレイから得られる差分値により部分係数を近似する必要があるが、サブアレイを構成する素子間隔を一定とした場合、次数  $R$  の値や音波の周波数に応じて、音圧の再現精度が異なることを確認した。

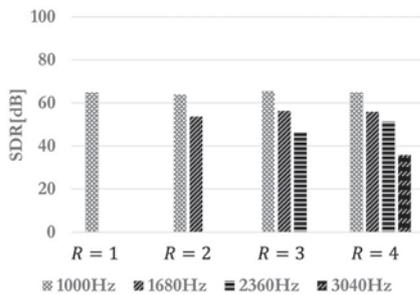


Fig.1 Reproduction accuracy with difference approximation.

## 2-1P-14

### 2-1P-14 パラメトリックスピーカによる反射音のサラウンド応用に関する検討

A study on the application of reflected sound by a parametric speaker to surround sound.

☆加藤弘士朗, 上原正志, 大川茂樹(千葉工大), 福島学(日本文理大)

- ◆ パラメトリックスピーカ(Parametric Array Loudspeaker: PAL)による指向性の鋭い放射音を受聴者の後方で反射させ、サラウンドシステムにおける後方スピーカとして利用。
- ◆ PAL のビーム幅やビームの減衰に関する反射の有無からの検討。
- ◆ 反射ありが反射なしと比較してビームの距離減衰が大きい。
- ◆ 反射なしでは局所的な大きな振幅が見られたが、反射ありでは見られなかった。

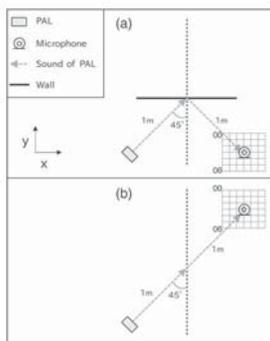


Fig.1: Top view during the experiment

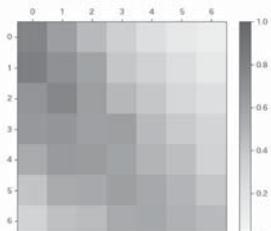


Fig.2: Heatmap of the maximum value of impulse in the condition with reflection

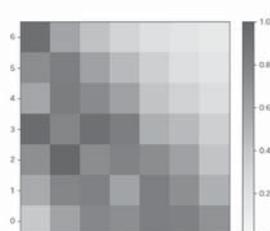


Fig.3: Heatmap of the maximum value of impulse in the condition without reflection

## 2-1P-13

### 2-1P-13 ウェアラブルネックスピーカと頭部伝達関数を用いた3次元音像定位

Three-dimensional sound image localization using wearable neck-speaker and head-related transfer functions

☆柿崎洋哉(千葉工大・院), 飯田一博(千葉工大・先進工)

- ◆ 5 機種(機種)のネックスピーカと本人の頭部伝達関数を用いて、水平面 12 方向および正中面 7 方向を目標方向とした音像定位実験を行った。ただし、ネックスピーカの伝達関数の補正は行っていない。
- ◆ 水平面では、機種によって平均定位誤差が大きく異なり 11.7°-51.2°であった。
- ◆ 正中面では、全体的に平均定位誤差が大きく、32.0°-118.4°であった。

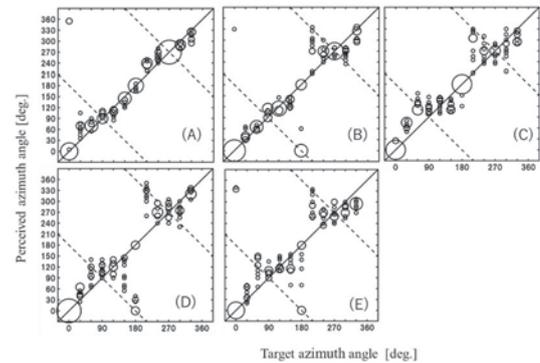


Fig.1 Localization responses to the stimuli in the horizontal plane for subject OKM using five kinds of wearable neck-speakers ((A)-(E)).

## 2-1P-15

### 2-1P-15 フレキシブル超指向性スピーカを用いた可変オーディオスポットの検討

Towards a Variable Audio Spot Using Flexible Parametric Array Loudspeakers

○宮川剛志, 近藤和弘(山形大院・理工学研)

- ◆ フレキシブル超指向性スピーカを用いることでオーディオスポットの範囲を対象に合わせて変化させる。
- ◆ Kinect を用いて複数人の頭部を検出し、頭部の中心座標を算出することで音声提示すべき範囲を最低限にする。
- ◆ PC で Kinect の映像を録画する。この映像を解析することで中心座標の追尾精度を測定する。
- ◆ 被験者 2 人を計 6 パターンの位置で移動させ中心座標の追尾精度を測定した。
- ◆ 中心座標の測定精度は最も低い値であっても 80%以上の精度を確認した。
- ◆ 対象の頭部座標の間隔が小さいほど中心座標の追尾精度が高くなることを確認した。

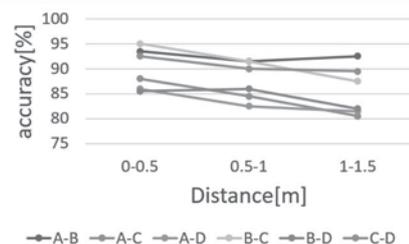


Fig.1: Relationship between center coordinate tracking rate and distance.

### 2-1P-16

#### 2-1P-16 下方に配置した2つのスピーカによる上方音像の再生

On reproduction of elevated sound image by two loudspeakers placed below  
☆原田雄典, 羽田陽一(電通大)

- ◆下方に配置した2つのスピーカを用いて、受聴者の正面よりも上方に音像を呈示することを目標として、スピーカ間隔と高さの異なる複数セットの中の2つのスピーカを用いて音を再生し、上下方向の音像定位実験を行った。
- ◆仮想音源位置を上方の1箇所に固定し、再生スピーカの位置を変えて行った受聴実験では、やや下方に配置した2つのスピーカを用い、狭帯域にしてトランスオーラルシステムを適用することで、上方音像を呈示できることが分かった。
- ◆再生スピーカの位置を2箇所に固定し、仮想音源位置を変えた受聴実験では、周波数帯域と再生スピーカの高さによって音像の高さに変化することが分かった。ただし、狭帯域では仮想音源位置に関わらず上方のほぼ同じ位置に定位し、広帯域であっても上下方向の音像を制御することは難しい結果となった。



Fig.1: Experimental setup

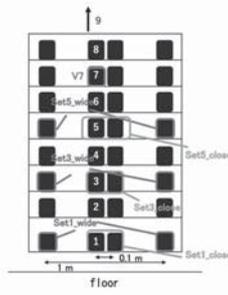


Fig.2: Schematic of experimental setup

### 2-1P-18

#### 2-1P-18 MMG 計測筋繊維活動推定量とエコー画像変位量の相関分析による対応の一検討

An investigation on Correlation Analysis between MMG-measured Muscle Fiber Activity Estimates and Echo Image Displacement

☆大里一矢, 山下涼介(NBU), 石川直生, 上原正志(千葉工業大学),  
△樋口幸(大分県立看護科学大), △市田秀樹(信州大), 福島学(NBU),  
△石井秀樹, △森竹隆広(シェルエレクトロニクス(株)), 柳川博文(arsl)

- ◆比較的容易に計測可能な上腕部に着目
  - ・上腕部の曲げ伸ばし(2秒間隔)を計測(Fig.1)
- ◆エコー基準(曲げ伸ばし1回分)でMMG(曲げ伸ばし3回分)で評価
- ◆エコーを加速度に変換(Fig.3)しMMGと相関分析(Fig.5)
  - ・大きな動き(曲げ伸ばし動作)に相関が見られた(Fig.5(c))
  - ・2500 sample・7200 sample・10800sample 付近に正の相関

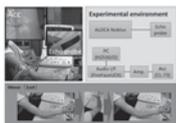


Fig.1 Experiment environment

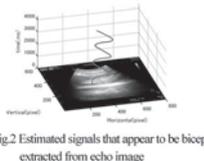


Fig.2 Estimated signals that appear to be biceps extracted from echo image

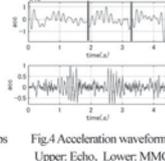


Fig.4 Acceleration waveform Upper: Echo, Lower: MMG

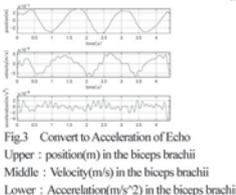
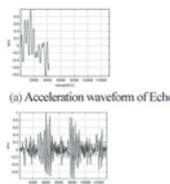


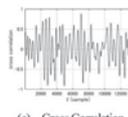
Fig.3 Convert to Acceleration of Echo Upper: position(m) in the biceps brachii Middle: Velocity(m/s) in the biceps brachii Lower: Acceleration(m/s^2) in the biceps brachii



(a) Acceleration waveform of Echo



(b) Acceleration waveform of MMG



(c) Cross Correlation

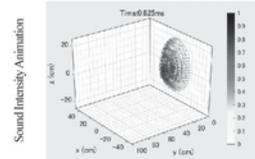
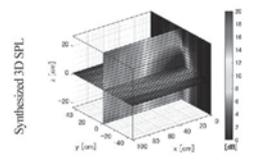
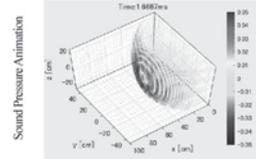
Fig.5 Result of Correlation Analysis

### 2-1P-17

#### 2-1P-17 3次元直方体音場の自動測定および可視化

Automatic Measurement and Visualisation over a Three-dimensional Cuboid Sound Field  
☆任逸, 羽田陽一(電通大)

- ◆音場再現や騒音制御などの研究では、実環境での3次元音場観測が課題となっている。
- ◆本報告では、2軸アクチュエータと直線マイクロホンアレイを用いた3次元音場の全自動測定システムについて述べる。また、測定した音場の可視化(3次元アニメーション)を試みた結果についても報告する。
- ◆測定装置と測定プログラムを用いて、バックロードホーン型ラウドスピーカの音場測定を行った。音場の2次元断面の合成による3次元表現、3次元音場のアニメーションおよび3次元近似インテンシティ場のアニメーションを確認した。



### 2-1Q-1

#### 2-1Q-1 対称調波畳み込み

Symmetric Harmonic Convolution

☆竹内博俊(東大院・情理), 大石康智, 柏野邦夫(NTT), 猿渡洋(東大院・情理)  
調波畳み込みはCNNで調波構造を捉える畳み込み手法。→目的周波数未満の周波数の情報を取り込めず、処理精度が低下する問題がある。

- 従来法: 混合調波畳み込み (Anchor Mixing)
  - 複数の基本周波数の調波畳み込みを混合するアプローチ。計算の冗長性が高く、時間・メモリが増大し使用しにくい。定式化の解釈が難しい。
- 提案法: 対称調波畳み込み (Symmetric Harmonic Convolution)
  - 調波成分と自身が調波となる成分を畳み込むアプローチ。計算の冗長性が少なく、時間・メモリ効率がよく使用しやすい。定式化の解釈が可能。
  - 計算時間は4倍、メモリ効率は5倍向上

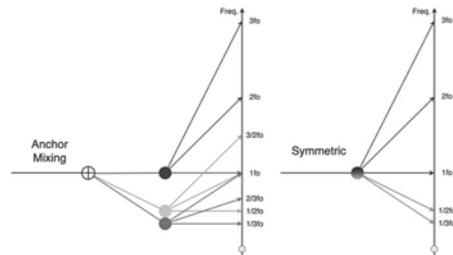


Fig.1: Conventional (left) and Proposed (right) methods. Each circle means a convolution, ⊕ means a mixture.

## 2-1Q-2

### 2-1Q-2 補助関数法に基づく 複数のチャンネル間時間差の同時推定

Multiple time delays estimation via auxiliary-function-based iterative updates

©山岡洗瑛, 小野順貴(都立大)

- ◆本研究では、複数のチャンネル間時間差の推定問題を扱い、補助関数法に基づき、これを同時推定する手法を提案する。
- ◆2チャンネル間の単一の時間差の推定法は、一般化相互相関(GCC)を用いた手法など、多数提案されている。
  - 複数の時間差推定問題では、参照マイクを設定し、各マイクとの時間差を順次求める必要があり、推定性能は参照マイクに依存する。
- ◆これまでに我々は、補助関数法に基づき、非凸関数であるGCCを目的関数とし、これを逐次的に最大化する手法を提案した。
- ◆本研究ではこれを拡張し、参照マイクを必要としない、複数の時間差の同時推定法を提案、高精度な時間差推定を達成した。

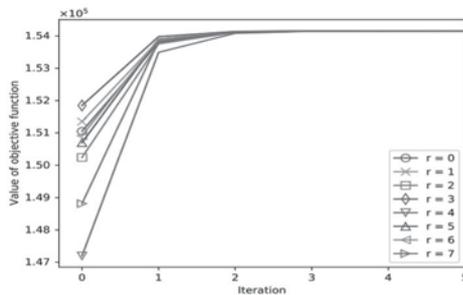


Fig. 1: Objective function via the proposed method, where index 0 corresponds to the initial value estimated by the conventional method with different reference microphone denoted as  $r$ .

## 2-1Q-4

### 2-1Q-4 非負値行列因子分解を用いた 音楽音響信号の音色変換

Tone conversion of music acoustic signals using non-negative matrix factorization

☆原口佳樹(静岡大・院・総合科学技術研), 竹内太法(静岡大学創造科学技術大学院), 立蔵洋介(静岡大・院・総合科学技術研)

- ◆近年の音楽編集の多くはコンピュータによって行われており、その基盤技術として音楽音響信号の信号処理は重要
- ◆非負値行列因子分解(Non-negative Matrix Factorization: NMF) [D. D. Lee+, 1999] を利用して、音源の演奏情報を保持したまま音色のみを別の楽器のものに置き換える音色変換手法を提案
  - 音階情報のような共通の情報のラベルによる、楽器間基底とアクティベーションの対応付け
- ◆単一音階の音源に対して基底を置き換える実験の結果、元の音源の演奏の情報を保持したまま音色のみを置き換えることに成功

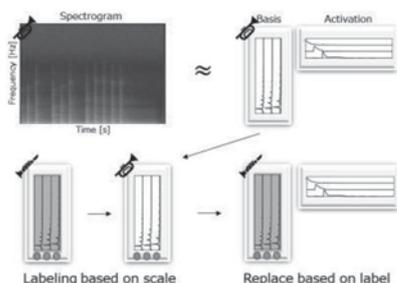


Fig. 1: Proposed Method

After decomposing the spectrogram, the basis is replaced based on the label. The color coding at the bottom of the basis vector represents the scale label.

## 2-1Q-3

### 2-1Q-3 パラメトリック/ダイナミックスピーカを用いたサラウンドシステムにおける 受聴者透過の移動音像構築

Moving Sound-image Construction Passing Listener in Surround System Using Parametric / Electrodynamic Loudspeakers

☆江川琢真, 中山雅人, 高橋徹(阪産大院)

- ◆我々はPAL (Parametric-Array Loudspeaker)とEDL (Electro-Dynamic Loudspeaker)を用いた音像距離制御に基づく新しい立体音響システムを提案している。この従来手法では、受聴者を透過させるような受聴者近傍の制御が困難であった。そこで、本稿では、受聴者の近傍と遠方で異なる処理を適用することで、音像距離制御に基づき、受聴者を透過する移動音像を構築する手法を提案する。
- ◆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。実音源(手でスピーカを移動させた条件)、従来手法、提案手法で構築した音像の移動感と頭内定位感を5段階評価のMOSで評価した。Fig.1に移動感、Fig.2に頭内定位感の実験結果を示す。Fig.1, 2において、Real(White)は手でスピーカを移動(ホワイトノイズ)、CM(White)は従来手法(ホワイトノイズ)、PM(White)は提案手法(ホワイトノイズ)の結果を示す。Fig.1よりCMに対し、Real, PMが両音源ともに滑らかな移動感であり、Fig.2よりReal, CMに対し、PMが、頭内に定位(受聴者を透過)しているから、提案手法の有効性が確認できた。

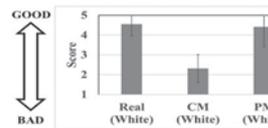


Fig. 1: Experimental result for moving sensation

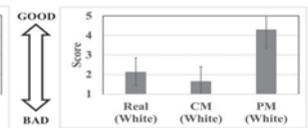


Fig. 2: Experimental result for intrahead sound localization

## 2-1Q-5

### 2-1Q-5 CNNを用いた船舶通知システムに向けた 航行音による騒音船舶識別方法の検討

A study on recognition method of noisy boat by boat noise using convolutional neural network for boat notification system

☆山口晴己, 武藤憲司(芝浦工大), 小林洋介(室蘭工大)

- ◆本研究の目的は音と視覚情報の主観評価を利用した運河沿い住民への船舶騒音印象軽減のための船舶接近通知システム開発である。
- ◆本報告ではシステムを構築する上で騒音と感じる船舶のみを事前に通知する手法の確立のため、CNNモデルに航行音の最大騒音レベルが85 dBを超えたものを騒音と感じる航行音としてラベルを付与し、騒音と感ぜない船のラベルとともに学習を行った。
- ◆CNNモデルの学習評価を行った結果、85 dB以上の航行音が生じた航行音に対し、6秒から60秒前に騒音と感ぜる航行音が到来することを予測することができた。

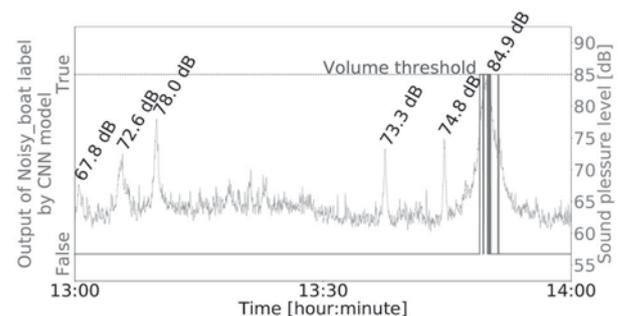


Fig. 1: Result of Noisy\_boat class output by CNN model with A-weighted sound level at the canal.

### 2-1Q-6

#### 2-1Q-6 直交検波を用いた音光変換デバイス「ブリンキー」のLEDの位置推定

LED position estimation of sound-to-light conversion device "Blinky" using orthogonal detection

☆佐藤直哉, 木下裕磨, 若林佑幸, 小野順貴(都立大)

- ◆我々は、マイクロホンとLEDにより音信号を光信号に変換できるデバイス「ブリンキー」を開発し、これを用いた新しい音響信号処理の枠組みを研究している。ブリンキーを用いた音響信号処理においては、ブリンキーのLED位置を特定し、発信される光信号を抽出する必要がある。しかしながら、多数のブリンキーを配置した場合に、人手でLED位置を特定することは大きな手間がかかる。
- ◆本研究では、ブリンキーが発信する光強度信号にパイロット信号を重畳し、自動でブリンキーのLED位置特定と信号の取得を行う手法を提案する。パイロット信号の成分を持つ画素を直交検波により検出することで、LEDの位置を推定する。
- ◆推定したLEDの位置をもとに、ブリンキーの信号の取得、およびパイロット信号を含まない信号への復元をする。
- ◆提案法により、Fig.1のようにLEDの位置が正しく推定できること、および信号の取得・復元ができることを、実環境実験により確認した。

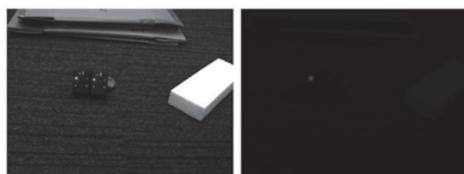


Fig. 1: One frame of the video used in the experiment (left), the result of orthogonal detection (right).

### 2-1Q-8

#### 2-1Q-8 DNNを用いた音声からの残響時間及びD/R推定

Joint Estimation of Reverberation Time and D/R from speech using DNN

☆宮田涼司, 羽田陽一(電通大)

- ◆残響時間及び直接音間接音比(直間比)は部屋の音響特性を表す重要な特徴量であり、音声認識などのアプリケーションにとって有用である。
- ◆CNNを用いたマルチタスク学習によって音声から残響時間、直間比に加え音源距離を同時推定することを検討した。
- ◆入力特徴量にはサイズの低減を目的としてメルスペクトログラムを使用した。
- ◆推定の結果、提案手法は計算機上で生成した1chの音声データより残響時間の推定が可能であることが示された。

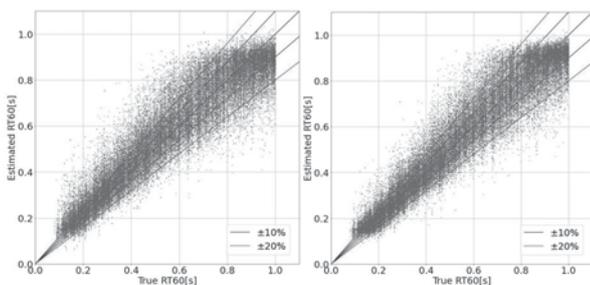


Fig. 1: Estimation results of reverberation time. (Left: multi task learning, right: single task learning)

### 2-1Q-7

#### 2-1Q-7 音光変換デバイス「ブリンキー」を用いた音響信号処理のための信号伝搬シミュレータ

Signal propagation simulator for acoustic signal processing with sound-light conversion device "Blinky"

☆岩本基裕, 木下裕磨, 若林佑幸, 小野順貴(都立大)

- ◆我々は、音光変換デバイスの「ブリンキー」を開発し、これとビデオカメラを用いて、新しい音響センシングの枠組みを研究している。しかしながら、ブリンキーを用いた音響信号処理手法を研究開発する際には、実環境にブリンキーを配置し、開発に必要なデータを収集しなければならない。このことは、ブリンキーの研究開発を難しくする一つの要因となっている。
- ◆本研究では、ブリンキーを用いた音響センシングにおける信号伝搬のシミュレータの開発を目指している。特に、この信号伝搬における光信号伝搬をモデル化、構築したモデルに従い光信号伝搬シミュレータを開発する。
- ◆収録した音声から、ブリンキーをカメラで撮影して得られる信号を復元する。
- ◆提案法により、Fig.1のように構築したモデルの妥当性、およびシミュレータの精度を実環境実験により確認した。

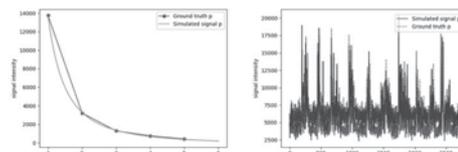


Fig.1: Optical signal propagation model (left), the result of simulator accuracy experiment (right).

### 2-1Q-9

#### 2-1Q-9 頭部包囲型球面マイクロホンアレイを用いた発話放射特性の分析について

Study on Speech Radiation Pattern by Head Enclosure Type Spherical Microphone Array

☆本地瑛, 羽田陽一(電通大)

- ◆よりリアルな遠隔コミュニケーションの実現のためには話者の発話放射特性の分析が必要である。本報告では母音の発話放射特性の分析を行い、話者性と音韻による違いについて調べた。
- ◆頭部包囲型48ch球面マイクロホンアレイを作成し、この装置とクロススペクトル法によって話者の頭部周辺の伝達特性を推定した。また、球面調和関数を利用して頭部周辺の立体的な発話放射特性を表示し、これについて分析を行った。
- ◆発話放射特性の立体的な表示により、音韻によって放射特性が異なることが確認できた。また、球面調和関数展開係数の実数表示により、母音 /i/ は母音 /a/ に比べて下方方向に強く放射していることが確認できた。

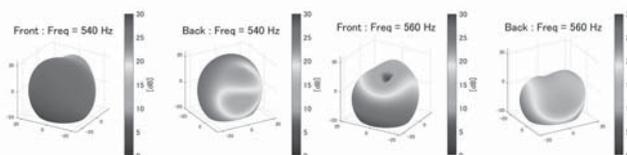


Fig.1: Speech Radiation Pattern (vowel : /a/)

Fig.2: Speech Radiation Pattern (vowel : /i/)

2-1Q-10

2-1Q-10 風雑音スペクトルのクラスタリングに基づく風向風速計の検討

Examination of anemometer based on wind noise spectrum clustering

☆近藤正和(静岡大学大学院総合科学技術研究科), 竹内太法(静岡大学創造科学技術大学院), 立蔵洋介(静岡大学大学院総合科学技術研究科)

- ◆目的  
マイクロホンアレイを用いて風雑音から風向風速を計測
- ◆提案手法  
各風向風速の風雑音を収録し,SVM を用いてクラスタ分類  
入力された風雑音がどのクラスタに属するかを判定して識別
- ◆結果  
3チャンネルのマイクロホンアレイを用いて風向の識別実験を行った結果,隣接していない風向の誤識別が少ないことを確認  
1つのチャンネルにおいて風向がマイクの前か後かかの判定が可能であることを示唆

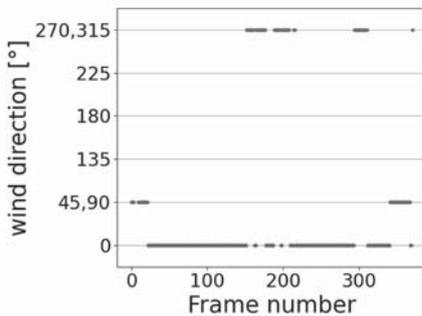


Fig. 1: Wind direction identification result for wind noise test data in the 0 degree direction.

2-1Q-12

音場補間に基づく  
円状マイクロフォンアレイの自己回転角度推定  
A study of estimating the rotation angle  
of a circular microphone array

☆連冠三, 中嶋大志, 若林佑幸, 小野順貴(都立大)

- ◆アレイ信号処理は音響伝達系が時不変性と仮定しているが,実応用を考えるとその仮定は満たされず,系の変化により音源分離や音源強調などの空間フィルタの再計算が必要となり,実時間処理が困難となる.
- ◆先行研究において,円状アレイの回転移動による系の変化に対応するための音場補間手法が提案された.しかし,音場補間のために回転角度が既知であることが要求されている.
- ◆本研究では,音響情報のみから円状アレイ自己回転角度の推定手法を提案する. Fig. 1 に提案法の手順を示す.観測ベクトルが多次元複素正規分布に従うことを仮定し,上記音場補間による推定観測ベクトルを利用した最尤推定から角度を推定する.
- ◆推定性能確認のため計算機シミュレーションを行った. Fig. 2 のように尤度最大となる角度が真の回転角度であることを確認した.

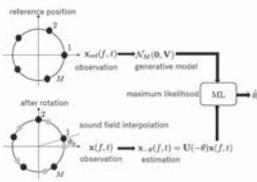


Fig. 1: The block diagram of estimating the rotation angle of a circular microphone.

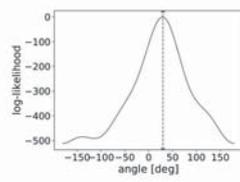


Fig. 2: The Log-likelihood for one source with the rotation angle of 30 deg.

2-1Q-11

2-1Q-11 DNN と CAE を用いた構造物内部の音源探査に関する研究  
-Autoencoder を用いた実環境のデータ変換の適用-

Study on sound source localization inside a structure using DNN and CAE. -Applying real-world data transformation with autoencoder-

◎喜多俊輔(大阪技術研), 梶川嘉延(関西大)

- ◆Deep Neural Network (DNN)と Computer Aided Engineering(CAE)を利用した構造物内部の音源探査手法を提案し,シミュレーションおよび実験ドメインのそれぞれにおいて,音源探査が可能であることを報告してきた.
- ◆しかし,シミュレーション上で DNN により構築した音源探査モデルを実環境に適用した際,音源探査精度は著しく悪化する課題があった.
- ◆この課題解決を目的に,シミュレーションと実験のデータセットの一部を用いて,実験データからシミュレーションデータへ変換する Autoencoder(AE)を導入することで,音源探査精度が向上するか検証した.
- ◆AE によるデータ変換および,周波数応答をマスキングする Data augmentation (DA)を利用することにより,構造物内部の音源に対する音源探査の精度が向上することがわかった(Table 1).

Table 1 Results.

Index	(1)AE	(2)AE-DA	(3)SID-DA	Learning (RMSE [mm])	Test (RMSE [mm])
(I)	×	×	×	92.98	141.94
(II)	×	×	○	28.29	142.30
(III)	○	×	×	92.98	109.29
(IV)	○	○	×	90.19	104.72
(V)	○	×	○	27.50	124.83
(VI)	○	○	○	26.93	76.16

2-1Q-13

2-1Q-13 音の到来方向推定と機械学習による音源位置と種類の把握方法  
-カラスの鳴き声を対象としたリアルタイム識別の試み-

Identification method of position and kind of the sound sources by the estimation of arrival direction of sound and the machine learning  
-An attempt to identify crow calls in real time-

○石井要次, 土肥哲也(小林理研)

- ◆安価なマイクロホンアレイおよび機械学習を使用して,遠隔地でリアルタイムに音源位置および音源種類を把握できるシステムを開発した.
- ◆本システムは,エッジコンピューティングおよびIoTの概念を組み合わせられており,各ユニットで求めた音の到来方向情報やスペクトログラムのみを転送することでデータ転送量を削減している.
- ◆カラスが鳴いた時に各アレイが推定した音源方向の例を Fig.1 に示す.すべてのアレイの結果は,図中★印方向を示しており,人間が確認したカラスの位置と一致した.
- ◆さらに,スペクトログラム情報に基づいた機械学習によってカラスの鳴き声を自動識別機能も搭載し,SN 比が低い場合を除いてカラスの鳴き声を自動識別できることを確認した.

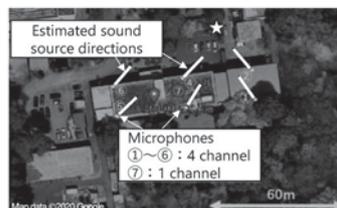


Fig.1: Result of the verification experiment on the estimation of the sound source position: Lines indicate that the estimated sound source directions at each recording point.

### 2-1Q-14

#### 2-1Q-14 拡散性雑音をモデル化した独立低ランク行列分析

On blind source separation in diffuse noise fields

◎池下 林太郎, 伊藤 信貴, 中谷 智広 (NTT)

○ 拡散性雑音環境における音源分離問題(周波数  $f$ , 時間フレーム  $t$ )

$$x(f, t) = \sum_{i=1}^K a_i(f) s_i(f, t) + n(f, t) \in \mathbb{C}^M$$

- ・  $K$ : 音源数,  $M$ : マイク数,  $K \leq M$  と仮定
- ・  $a_i(f) \in \mathbb{C}^M$ : ステアリングベクトル for 音源  $i = 1, \dots, K$
- ・  $n(f, t) \sim \mathcal{CN}(0_M, \lambda(f, t)V(f))$ : 拡散性雑音
- ・  $V(f) \in S_{++}^M$ : 空間相関行列 (正定値)

○ 研究の動機

- ・ MNMF [Ozerov+2009] に対する EM/MM アルゴリズムは収束が遅い
- ・ ILRMA [Kitamura+2016] は拡散性雑音をモデル化していない

○ 本研究のアイデア: MNMF を分離モデルに変数変換する(近似は詳々)

○ 変数変換を与える命題 ( $K \leq M$  と仮定)

$A_1 = [a_1, \dots, a_K] \in \mathbb{C}^{M \times K}$  と  $V \in S_{++}^M$  に対して,  
 $W = [w_1, \dots, w_M] \in \mathbb{C}^{M \times M}$  と  $G \in S_{++}^K$  が存在して, 次を満たす:  
 ・  $W^h [a_1, \dots, a_K] = [e_1, \dots, e_K] \in \mathbb{C}^{M \times K}$  ( $W$  は音源分離を達成する)  
 ・  $W^h V W = \begin{bmatrix} G & O_{K, M-K} \\ O_{M-K, K} & I_{M-K} \end{bmatrix}$  (雑音の空間相関を  $G$  まで無相関化する)  
 特に, 各  $w_i$  ( $i = 1, \dots, K$ ) は LCMV ビームフォーマとなる:  

$$w_i = V^{-1} A_1 (A_1^h V^{-1} A_1)^{-1} \in \mathbb{C}^M$$

○ 提案手法: NoisyLRMA = MNMF + 命題 +  $G$  を対角と近似

- ・ NoisyLRMA はランク制約付き FastMNMF [Sekiguchi+2020] と等価

### 2-1Q-16

#### 2-1Q-16 雑音スペクトルの振幅・位相同時推定に基づくスペクトラルサブトラクションによる雑音抑圧

Noise suppression by spectral subtraction based on simultaneous estimation of amplitude and phase of noise spectrum

☆竹中幸輝, 小澤賢司(山梨大・院・医工農学総教)

◆ マイクロホンアレイの観測信号を画像とみなして2D FFTすることにより2次元スペクトルを得る。その空間方向の直流成分には、(1) アレイに直交する方向 ( $0^\circ$ ) から到来した目的音の全スペクトルと、(2) その他の方向から到来した雑音の遅延和アレイ出力に相当する成分が混在している。ニューラルネットワーク (以下 NN と略記) により、(2) の成分を推定し減算することで雑音抑圧を行う。

◆ 前報では、雑音スペクトルの振幅と位相を独立した NN で推定していたため、位相特性における振幅特性への依存が反映されていなかった。

◆ 振幅と位相を同時に推定可能な NN モデルを使用することにより、位相の学習精度が向上した。また、2つの白色雑音に対する抑圧量については前報と同程度の性能を示した (Fig. 1)。

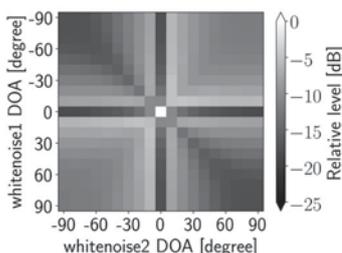


Fig. 1: Noise suppression for two noise sources.

### 2-1Q-15

#### 2-1Q-15 音源クラス識別器学習を用いた一般化多チャンネル変分自己符号化器法

Generalized Multichannel Variational Autoencoder with Auxiliary Source Label Classifier Training

◎関 翔梧, 亀岡 弘和(NTT)

◆ 劣決定条件下での多チャンネル音源分離問題において、これまでに音源クラスを補助情報とする条件つき変分自己符号化器 (CVAE) を用いた多チャンネル変分自己符号化器法 (GMVAE) が提案されている。

◆ 音源信号のパワースペクトログラムのモデル化に CVAE を利用する有効性が示されている一方で、CVAE には補助情報に関する制約がないため、条件づけが低下し音源クラスが無視される可能性がある。

◆ 補助情報に関する条件づけの影響を強調するために、音源クラス識別器つき変分自己符号化器 (ACVAE) を音源モデルとする GMVAE 法を提案し、音源クラス識別器の学習方法による影響を調査する。

◆ 実験的評価より、提案法による分離性能の改善を確認した。

Table 1: Separation performances (a) Speaker-closed condition

Method	SDR	ISR	SIR	SAR
CVAE	4.1	9.8	7.5	8.4
ACVAE (Intra)	4.0	9.8	7.4	8.4
ACVAE (Inter)	4.4	9.9	7.6	<b>9.1</b>
ACVAE (Intra & Inter)	<b>4.5</b>	<b>10.1</b>	<b>7.9</b>	8.8

(b) Speaker-open condition

Method	SDR	ISR	SIR	SAR
CVAE	3.4	9.2	6.6	7.7
ACVAE (Intra)	3.5	9.5	6.7	7.7
ACVAE (Inter)	<b>3.7</b>	9.7	6.7	7.9
ACVAE (Intra & Inter)	<b>3.7</b>	<b>9.8</b>	<b>7.0</b>	<b>8.1</b>

### 2-1Q-17

#### 2-1Q-17 ペアデータを必要としない敵対的学習に基づく多チャンネル音源分離

Multi-channel speech source separation based on adversarial learning without need for paired data

☆中込 優(早大), 戸上 真人(LINE), 小川 哲司(早大), 小林 哲則(早大)

◆ ペアでない混合音声とクリーン音声を用いた敵対的学習に基づく音源分離法を提案

➢ **Generator**: 分離 DNN とその後段に MMF を持つ主分離器

➢ **Discriminator**: クリーン音声信号と Generator による分離音声信号の中間特徴量が離れるように学習する Auto-Encoder

Generator はクリーン音声らしい分離音を出力するように敵対的学習を行うことで、ペアデータを必要とせず音源分離法を学習可能 (Fig. 1)。

◆ 学習安定化のための2つの施策

① クリーン音声に対し Generator がもつフィルタ特有の歪みを付与し、DNN のパラメタの学習を音源の分離に注力させる。

② 観測音に対する分離音の尤度最大化により Generator を事前学習。

◆ 評価実験: 提案法の有効性を確認し、歪み付与と事前学習により性能改善を確認 (Table.1)。

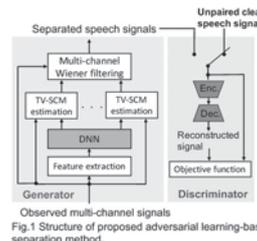


Table 1. Performance of existing and proposed speech source separation methods. The proposed method shown in the fifth row was pre-trained by unsupervised training [1].

	DNN training	SDR [dB]	PESQ
Unprocessed	-	0.80	1.69
Unsupervised training [1]	-	5.12	1.78
Prop.	scratch	4.53	1.74
Prop. w/ MWF dist.	scratch	4.64	1.74
Prop. w/ MWF dist.	Pretrain [1]	5.67	1.81

[1] L. Drude et al., in INTERSPEECH, 2019

2-1Q-18

2-1Q-18 メディアン型HPSSを用いた時間周波数マスクに基づくブラインド音源分離

Blind source separation based on time-frequency mask using median-type HPSS

☆大藪宗一郎, 北村大地 (香川高専), 矢田部浩平 (早稲田大)

時間周波数マスクに基づく優決定BSS (TFMBSS) は, 時間周波数マスクに基づいて線形の (歪みの少ない) 多チャンネル音源分離が可能である. この利点を活かして, 調波打撃音分離 (HPSS) に基づく時間周波数マスクをTFMBSSの音源モデルとして用いた手法を提案してきた.

本稿では, より良い音源モデルの探求のため, メディアン型HPSSに基づく時間周波数マスクをTFMBSSに新たに導入した手法を提案する. また, 提案手法におけるフィルタサイズやスムージングパラメータについて実験的に調査し考察する. さらに, 従来のHPSSに基づくTFMBSSと性能を比較し, 提案手法の有用性の検討を行う.

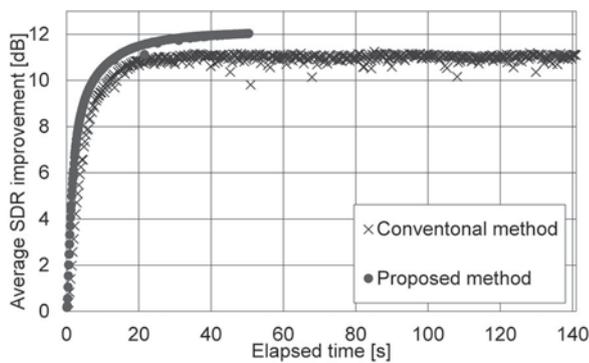


Fig. 1: Average convergence behaviors of SDR improvements in conventional and proposed methods in terms of elapsed time.

2-2P-2

2-2P-2 自己教師あり学習によるラベル無し自由発話を用いた構音障害者音声認識

Dysarthric speech recognition using unlabeled speech with self-supervised learning

☆澤佑哉, 富士原健斗 (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)

- ◆ 構音障害者の自由発話音声は発話内容を認識し文字起こしを行うことが困難であり, ラベルの無い音声データの活用方法が求められている.
- ◆ 本研究では, ラベルの無い構音障害者の自由発話音声を用いて自己教師あり学習を行い, 学習したモデルを音声認識に流用することで音声認識精度の向上を行う.
- ◆ 実験結果から, 自己教師あり学習を使用する場合に認識性能が向上し, 加えて自己教師あり学習における健常者音声を用いた事前学習が効果的であることが分かった.
- ◆ また, 自己教師あり学習に使用する発話数による認識性能の変化について調査し, 少ないデータ量でも認識性能が向上することが分かった.

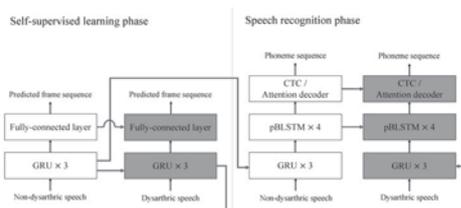


Fig. 1: System overview.

2-2P-1

2-2P-1 Estimation of Glottal Source Waveforms and Vocal Tract Shapes Based on ARMAX-LF Model

Kai Li (JAIST), Yongwei Li (CASIA), Jianwu Dang (JAIST), Masashi Unoki (JAIST), Masato Akagi (JAIST)

- ◆ This paper extended the system proposed recently using the ARX-LF model to autoregressive moving average exogenous (ARMAX)-Liljencrants-Fant (LF) straightly to obtain stable and accurate estimates of poles and zeros.
- ◆ The accuracy and stability of ARMAX-LF model were evaluated statistically in synthesized vowels and consonants by the average error rate (Table 1); and in natural speech by mean and standard deviation (Table 2). The results shown that our extended model can reach superior performance in synthesized speech and acceptable performance in natural speech /m/.

Table 1: The performance of ARMAX-WN, ARX-LF model and ARMAX-LF model in synthesized vowels and consonants

	Glottal source			Vocal tract		
	$E_p(\%)$	$E_z(\%)$	$E_{\omega}(\%)$	Pole( $\%$ )	Zero( $\%$ )	Average( $\%$ )
Vowels	ARMAX-WN	/	/	2.15	/	2.15
	ARX-LF	8.78	8.24	126.73	1.78	1.78
	ARMAX-LF	15.44	13.72	93.27	1.80	1.80
Consonants	ARMAX-WN	/	/	4.06	6.44	5.25
	ARX-LF	13.83	13.27	167.97	3.57	3.57
	ARMAX-LF	14.03	14.54	166.73	6.43	2.20

Table 2: The performance of ARMAX-LF model in natural speech /m/. M and SD refer to the mean and standard deviation of estimated values respectively.

Utterance	1	2	3	4	5	Mean						
M	SD	M	SD	M	SD	M	SD					
F1	232.9	101.06	227.0	103.32	224.2	8.85	217.4	6.24	220.1	6.29	224.3	8.35
F2	1493.4	51.10	1428.6	77.94	1493.7	70.26	1265.6	77.16	1210.8	79.74	1342.4	71.24
F3	2238.0	23.78	2233.6	79.31	2230.9	43.60	2118.0	77.78	2228.5	29.99	2209.8	49.09
F4	3181.2	27.94	3208.5	48.42	3185.5	50.35	3009.5	42.78	3135.0	76.72	3143.9	49.24
F5	4181.1	46.05	4224.6	65.20	4259.4	39.42	4212.6	72.18	4223.0	96.02	4223.5	64.17
Z1	679.7	302.26	747.6	207.91	738.2	233.23	733.5	130.03	674.9	321.96	678.6	219.08
Z2	2597.6	326.82	2544.2	317.46	2393.2	336.86	2381.6	357.03	2948.2	329.25	2573.0	339.48
Z3	3287.9	560.83	3289.5	456.43	3110.0	558.99	3141.4	425.53	3160.5	580.06	3197.9	516.37

2-2P-3

大規模日本語音声データを用いた end-to-end 音声認識の性能比較

A comparison of end-to-end models using large scale Japanese speech

○小島淳嗣 (アドバンスト・メディア)

- ◆ ストリーミング end-to-end 音声認識モデルである Monotonic Chunkwise Attention (MoChA), Streaming Transformer-Transducer (T-T), RNN-Transducer (RNN-T) の性能を比較
- ◆ 実際のサービスから収集された, 議会, 社内会議, 講演などの音声 2500 時間で学習
- ◆ 系列長と認識性能の関係についても調査

表 1: 認識率

手法	CER
MoChA	52.0
Streaming T-T	24.0
RNN-T	14.0

表 2: 入力系列長と認識率

手法	CER (short)	CER (long)
MoChA	43.4	69.0
Streaming T-T	23.8	24.3
RNN-T	13.7	14.5

### 2-2P-4

#### 2-2P-4 BLSTMを用いた end-to-end 音声認識に対する信頼度推定

BLSTM-based confidence estimation for end-to-end speech recognition

○小川厚徳, 俵直弘, 叶高朋, マーク・デルクロア (NTT)

- ◆本研究では報告例の少ない end-to-end (E2E) 音声認識に対する信頼度推定に取り組む。
- ◆Transformerに基づく最新の CTC/attention ハイブリッド E2E 音声認識システムを用いる。信頼度推定モデルとして, multilayer perceptron (MLP), bidirectional long short-term memory (BLSTM), Transformer に基づく 3 モデルを比較評価する。補助特徴量として, CTC スコア, attention スコア, LSTM 言語モデルスコア, これらの重み加算スコアを用い, これら 4 種類のスコアを比較評価する。
- ◆実験により, BLSTM の信頼度推定モデルとしての有効性と, CTC スコアの補助特徴量としての有効性を確認した。

Table 1. Comparison of the three confidence estimation models.

Model	EER ↓	AUC ↑	NCE ↑
MLP	14.3	0.923	0.364
<b>BLSTM</b>	<b>11.6</b>	<b>0.947</b>	<b>0.442</b>
Transformer	13.9	0.925	0.370

Table 2. Comparison of the four auxiliary features.

Features	EER ↓	AUC ↑	NCE ↑
Only char embeddings	30.2	0.762	0.103
<b>With CTC score</b>	<b>13.6</b>	<b>0.928</b>	<b>0.371</b>
With attention score	14.5	0.922	0.369
With LSTM score	21.5	0.861	0.227
With weight sum score	15.6	0.919	0.354

### 2-2P-6

#### 2-2P-6 超高齢者音声認識のための CycleGAN を用いたデータ拡張

Data Augmentation with CycleGAN for Super-Elderly Speech Recognition

☆下笠音輝(山梨大院), 西崎博光(山梨大院), 北岡教英(豊橋技科大)

- ◆超高齢者音声認識のためのデータ拡張
  - 音声変換に CycleGAN を利用
  - 若年層音声から超高齢者音声へ音声変換
  - 既存の音声コーパスに変換した音声セットを加えるデータ拡張による認識精度改善の有用性を調査
- ◆【実験 1】: 既存の音声と変換音声の認識性能評価
  - 音声変換: JNAS → EARS (超高齢者音声コーパス)
  - 音声変換によって超高齢者音声の認識性能に影響
    - ◇ JNAS: 認識性能向上
    - ◇ EARS: 認識性能低下
      - 音声復元によって生じるノイズが要因
- ◆【実験 2】: 既存の音声コーパスに変換音声した音声セット (conv.) を加えた学習データによる認識性能評価
  - JNAS, EARS の既存音声コーパスに対し JNAS (conv.) で拡張
  - 既存の音声コーパスを JNAS (conv.) でデータ拡張することにより認識性能向上
    - ◇ JNAS, EARS 共に認識性能向上
    - ◇ データ拡張としての有用性を確認

### 2-2P-5

#### 2-2P-5 連続発話手法および仮想環境を用いた車室内音声コーパス収集法の比較

Comparison of in-car speech collection methods using forced utterances and virtual environments

○水谷生風樹, 李晃伸(名古屋工業大学), 神沼充伸(日産自動車)

- ◆連続発話を用いた手法として連続発話想起法による効率的な対話コーパス収集法を提案, その収集効率, 収集したデータの質を評価
- ◆収録環境として実車環境, 仮想環境 (ドライブシミュレータ (DS)), オンライン環境による収録を比較
- ◆連続発話は通常の対話収録に比べて偏りのないコーパスが短時間で収集できた。また, 仮想環境やオンライン環境でも実車環境と同程度のコーパスが収集できた。

Table 1: Comparison experiment conclusion of methods and environments

	総実験時間(分)	総発話数	1発話あたりの実験時間(秒)	カバレッジ(%)
DS-対話	440	208	88.5	88.46
DS-連続	440	784	33.6	93.27
実車-連続	619	1033	36.0	99.03
オンライン-連続	213	309	41.4	84.61

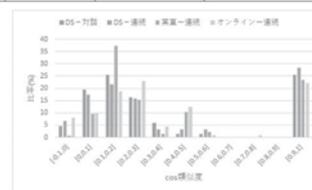


Fig. 1: Cos similarity of recorded words for each method and environment

### 2-2P-7

#### 2-2P-7 条件付きボルツマンマシンによる位相復元の初期検討

Initial study of phase recovery by conditional Boltzmann machine

☆羽賀洋克(電通大), 矢田部浩平(早稲田大), 岸田拓也, 中鹿亘(電通大)

- ◆Griffin-Lim 法は位相復元に用いられるアルゴリズムの一種だが, 初期値に大きく依存し, 不自然なアーティファクトを生成してしまう。
- ◆位相と振幅の関係を学習できるニューラルネットワークとして条件付き制限ボルツマンマシンを設計。可視層, 隠れ層, 制限項がそれぞれ完全二部グラフを構成しているモデルである。
- ◆既存の条件付き制限ボルツマンマシンでは可視層に Gauss 分布を仮定しており, 周期変数である位相には不適切だったため, von Mises 分布を仮定した条件付き方位ボルツマンマシンを提案。位相と振幅の学習が可能になった。
- ◆提案モデルで Griffin-Lim の初期値を推定した結果, 復元した音声の品質がわずかに上昇した (Fig.1)。

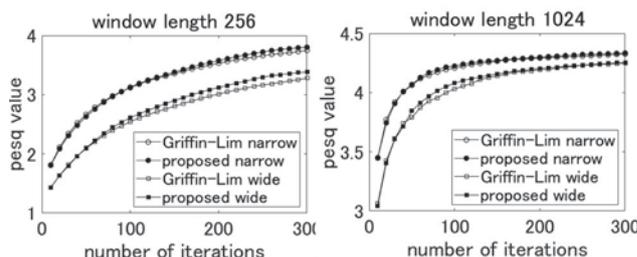


Fig1: Variation of psq values by the number of Griffin-Lim

### 2-2P-8

#### 2-2P-8 腕の動作を用いた母音と わたり音/y,w/のリアルタイム合成

A real-time synthesis of vowels and glides /y/ and /w/ using arm motions.

○伊藤仁, 小島銀河, 遠藤慎也(東北工業大学)

- ◆腕の動作でフォルマント周波数を逐次的に制御する音声合成システム (Fig.1) を用いて, わたり音/y/と/w/の知覚に必要な高速なフォルマント遷移をどの程度実現できるか検討した.
- ◆フォルマント遷移速度に応じて二連母音からわたり音節へと知覚が変化する 16 種類の Target 音を用いて, 16 名の被験者に動作でこれを再現させる音声合成実験を行った.
- ◆Target 音と動作による合成音声のフォルマント周波数の RMS 誤差は, 手の空間位置を用いた Position 制御で平均 363 cent, 掌の角度を用いた Direction 制御で 303 cent であった.
- ◆この制御法による再現精度の違いは, 目標とする軌跡に対する動作の空間的なずれではなく, 動作の時間的な特性, 特にオーバーシュート量に起因することが分かった (Fig.2).

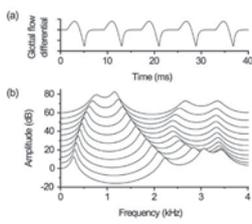


Fig.1: (a) source signal  
(b) Vocal-tract filters

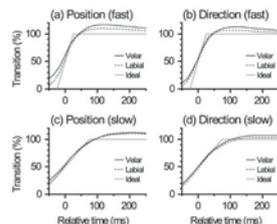


Fig.2: Normalized spectral transitions.

### 2-2P-10

#### 中国語の唇音の前鼻音と後鼻音の発話自動判別に関する研究

Study on Automatic Utterance Discrimination of Front Nose and Back Nose Consonants in Chinese Bilabial

○星野朱美(富山高専)

- ◆背景と目的: 日本語話者は中国語の「前鼻音」と「後鼻音」を聞き取りと発話が困難な上, 自習する際にも自分の発音に対する正確な評価手段がない. CAI のための発話自動学習システムの開発を試みた.
- ◆手法: 本研究では中心周波数が 50Hz~6850Hz, 帯域幅 200Hz の 35 チャンネルのフィルターバンクを用いた周波数スペクトル自動測定により, 日本人学生・中国語話者各々 10 名の発話を分析して, それぞれの鼻音の発話の有声期間中のパワーと F1~F3 の特性を抽出し, 「前鼻音」と「後鼻音」の特徴的パターンにより判定基準を求めた. 更にそれらのパターンを用いて「前鼻音」と「後鼻音」の発話の自動判別システムを開発した.
- ◆自動判別: 本システムを用いて, 日本人学生 10 名と中国語話者 5 名の発話の自動判別の実験を行った. Fig.1 は「前鼻音」と「後鼻音」の唇音音節の自動判別の流れ図である.
- ◆結果: 中国語の「前鼻音」と「後鼻音」の発話はそれぞれ 78%と 89% 以上の良好な自動判別率が得られた.

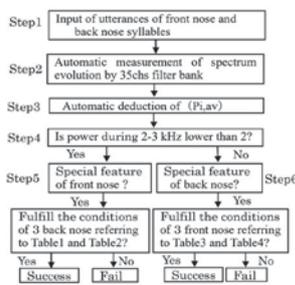


Fig.1 Discrimination diagram of front nose and back nose labial syllables

### 2-2P-9

#### 2-2P-9 「転声こえうらない」を通じたボイスチェンジャー 品質改善のための定性分析と考察

Consideration for Improving Voice Conversion Quality through noise classification of "Koedurana: Virtual Voice Incarnation"

☆堀部貴紀 (REALITY/明治大), △白井暁彦 (REALITY), 森勢将雅 (明治大)

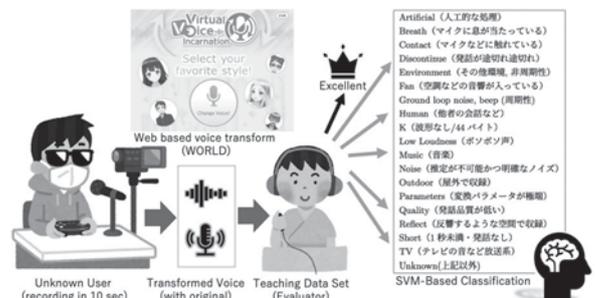


Fig.1: SVM-based classification from anonymous user voice samples.

ボイスチェンジャー目的の音声に混入するノイズを自動分類する手法について検討した. ノイズを含む音声に対して Fig. 1 右のノイズ要因に分類し, 教師データセットを構築した. 線形・非線形のモデルとそれぞれ多クラス分類に One-vs-One と One-vs-Rest の 2 つの手法を用いて計 4 モデルで学習を行った. 非線形モデルにおいて有効な分類性能を確認したが, 加えて音声テキスト変換 (Speech to Text) を前段に使用することで高速な一次分類が可能になるだけでなく, 言語的な内容についても評価できる設計が可能になったと考えられる.

### 2-2P-11

#### 2-2P-11 ポップアウトボイスの主観評価値と対応する声道形状に基づく指標に関する検討

A study on the indices based on vocal tract shape representing subjective score of pop-out voice

☆奥田康弘, 坂野秀樹(名城大院).

- ◆雑音環境下でも際立って目立つ音声をポップアウトボイスと呼ぶ. 本研究では, ポップアウトボイスにおけるポップアウトの程度と, 口唇や舌などの声道形状の動きとが関連するかどうかを調査する.
- ◆声道形状の動きを表す指標として, 音声データから声道断面積関数, 対数断面積比を抽出し, その時間差分に基づく指標を算出した.
- ◆ポップアウトの程度をあらゆる主観評価値と声道形状の変動をあらゆる指標との相関を調査した. その結果, 算出した相関係数が極めて小さい値となっており, 相関を確認することはできなかった.
- ◆原因としては, 抽出する声道形状の時間変動に不自然なゆらぎがあるため, 時間差分を用いた指標では適切に声道形状の動きを反映できていない点が考えられる.

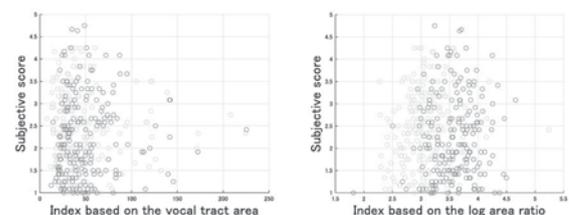


Fig.1 Scatter diagram of subjective score versus the indices. Left: the index based on the vocal tract area. Right: the index based on the log area ratio.

### 2-2P-12

#### 2-2P-12 ロングトーン母音歌唱を対象とした 基本周波数変化に伴う 声道断面積関数変化の統計的解析

Statistical analysis of relationship between vocal tract area function and fundamental frequency in long-tone vowel singing

☆金井郁也, 森勢将輝(明治大)

- ◆目的: 歌唱音声における基本周波数変化と音色変化のモデル化
  - エフェクターによる大幅な基本周波数補正で生じる個人性の破綻
  - 音色変化を表すうえで適切な特徴量は何か?
- ◆先行研究: 基本周波数変化に伴う声道断面積関数の変化
  - MRI、CT等の機器による計測は困難
  - SparkNG[河原]により、音響信号から声道断面積関数を求める
- ◆解析結果: 発声周波数上昇に伴い、口唇から1cmにおける声道断面積関数が大きくなる傾向があることを確認

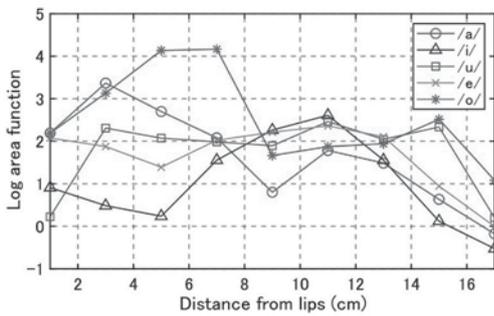


Fig. 1: Vocal tract area function in long-tone vowel singing

### 2-2Q-2

#### 2-2Q-2 単語分類タスクにおける時間領域特徴量と 周波数領域特徴量の組み合わせ

Combination of Time-Domain and Frequency-Domain Acoustic Features for Speech Commands Recognition

☆王イ康, 大川正暉, 西崎博光(山梨大院)

- ◆単語分類タスクにおける音響特徴量の組み合わせ手法を検討
  - 複数の分類モデルを用意し、異なる種類(時間領域, 周波数領域)の音響特徴量を各モデルに入力し、分類結果を事後確率ベースで結合し、最終結果を判定
- ◆実験
  - 各特徴量の最適モデル構造の考察
  - 時間領域特徴量と周波数領域特徴量のバッチ正規化必要性の考察
    - ◇ 時間領域特徴量はバッチ正規化を行わないほうがよい
  - 単語分類タスクで特徴量の組み合わせ手法を評価(表1)
    - ◇ 分類性能が向上: 単一特徴量と比較して1.45pt 向上
    - ◇ 雑音頑健性向上: 雑音音声に対してより頑健に

Table 1. Classification accuracies [%] of different features and combinations for various noise.

Features	Noise			
	no	gaussian	background	pink
raw waveforms	94.19	84.20	80.47	87.19
bit sequences	93.73	82.65	78.11	86.74
MFCCs	93.55	87.47	82.29	87.01
logfbank	94.01	88.92	82.47	87.92
bit + MFCCs	95.10	88.19	84.65	89.56
bit + logfbank	<b>95.64</b>	88.56	84.11	<b>90.28</b>
bit + raw waveform	94.37	85.38	82.47	87.83
raw waveform + MFCCs	95.46	88.20	<b>84.92</b>	88.10
raw waveform + logfbank	95.55	<b>89.92</b>	84.20	89.56

### 2-2Q-1

#### WFST を用いた医療音声認識における 専門用語へのバイアシングの検討

Study on biasing medical term using WFST for medical speech recognition

○小島淳嗣(アドバンスト・メディア)

- ◆医療音声認識において、薬品や病名などの専門用語は正確に認識されなくてはならない
- ◆end-to-end 音声認識において、固有名詞のリスト(バイアス語)から Weighted Finite-State Transducer (WFST)を構築し、shallow-fusion する手法が提案されているが、バイアス語が多い時(3000-5000語)、認識率が低下することが報告されている
- ◆診療科ごとにWFSTを構築し、end-to-end 音声認識モデルの予測した診療科に基づいて、WFSTを切り替える手法を検討

- ✓ 5000語以上の医療専門用語をバイアス語として使用
- ✓ バイアス語が多い時、認識率が低下する問題を回避

表1: 結果

手法	文字誤り率
バイアスなし	14.4
診療科+バイアス区間推定	8.8
+ 診療科非依存 WFST	15.4
+ 診療科依存 WFST	<b>5.8</b>

### 2-2Q-3

#### 2-2Q-3 音声認識深層ニューラルネットワークに とって重要なユニットの抽出と計算量削減

Extraction of important units for speech recognition deep neural networks and computational reduction

○金寺 登(石川高専)

- ◆ネットワーク内部でどのユニットがどの程度判別に寄与しているかが不明である。
- ◆そこで、本研究では深層ニューラルネットワーク内部の処理を可視化する方法として、ネットワーク出力を各ユニットで偏微分した値の変動(標準偏差)によって、識別の根拠となる重要なユニットを特定する方法を提案する。
- ◆音声認識深層ニューラルネットワークに適用し、その有効性と計算量低減効果を確認した。

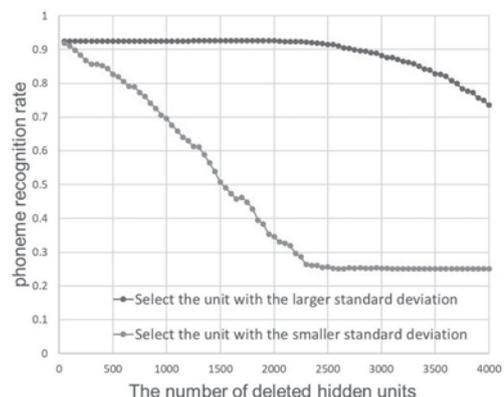


Fig. 1: Reduce computational costs by identifying critical units.

## 2-2Q-4

### 2-2Q-4 異種・複数の深層学習モデルを用いた音声 中の検索語検出方式の高精度・低メモリ化

Improving Retrieval Accuracy and Required Memory for Query-by-example  
Using Heterogeneous and Multiple Deep Learning Models

☆西野将弘, 小嶋和徳(岩手県立大学), 李時旭(産総研),  
伊藤慶明(岩手県立大学)

- ◆ SQ-STD: Posteriorgram 照合一検索精度○, 検索時間×, メモリ量×  
音声データ最尤系列化方式一検索精度×, 検索時間○, メモリ量○
- ◆ 提案手法: 深層学習モデルの出力ノードが character, syllable, triphone, monophone に対応する異種モデルを用い最尤系列化方式を実行。得られたスコアを線形和統合。最良統合割合は NTCIR-10(0.1-0.2-0.2-0.5), NTCIR-12(0.4-0.3-0.1-0.2)で最高精度。この割合で交差検証を実施
- ◆ Posteriorgram 照合との比較  
ONTCIR-10: Posteriorgram(BLSTM)から交差検証は+5.16pt (最良割合+5.39pt), 2.11秒(-26.97秒), 0.43GB(-114.21GB)  
ONTCIR-12: Posteriorgram(Hybrid CTC/Attention・syllable)から交差検証は+5.00pt (最良割合+5.06pt), 1.91秒(+0.19秒), 0.40GB(-1.91GB)

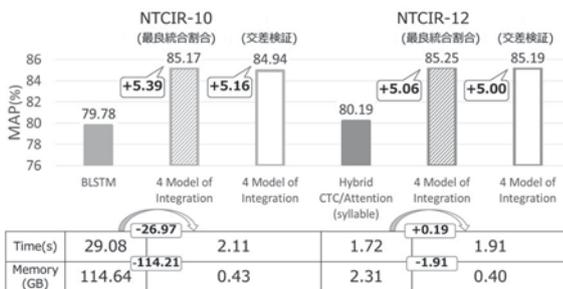


Fig. 1: Retrieval performance of the proposed method in NTCIR-10 and NTCIR-12.

## 2-2Q-6

### 2-2Q-6 日本語母語話者による日本語および英語 の流音の発音比較—予備的分析—

A preliminary comparison of liquid production in Japanese and English  
by Japanese speakers of English

○勅使河原三保子(駒澤大)

- ◆ 日本語母語話者は母語に流音が1種類しかないため、流音が2種類以上ある言語の流音を習得するのに苦労するのはよく知られているが、同一話者の日英両言語における流音の発音を比較した研究は知られていない。
- ◆ 本稿では10名の日本語母語話者が日英両言語で八つの質問を音読した後に質問に回答した音声に含まれたラ行子音(728個)と、/r/, /v音(各々365個、393個)、計1,486個を聴覚的に分析した。
- ◆ 英語の流音の生成パターンに基づき10名の話者を、/r/優位話者(/r/は概ね標準的だが、/v/を/r/で代用する)、/v/優位話者(/v/は概ね標準的だが/r/を/v/で代用する)、あまり発音に熟達していない話者(/r, v双方に代用がある)に分類して結果を検討した。
- ◆ /r/優位話者と/v/優位話者は日本語の流音の発音パターンにおいても異なり、/v/優位話者の日本語の流音は側音性(laterality)のある発音が最も多かった。

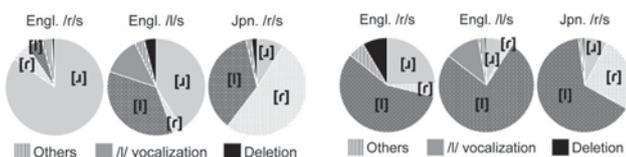


Fig. 1: Breakdown of pronunciations averaging across /r/-dominant speakers (left) and /v/-dominant speakers (right).

## 2-2Q-5

### テキスト音声合成及び遠隔発話特徴付与による学習 データ拡張手法を用いた End-to-End 音声認識

End-to-End Speech Recognition Using Data Augmentation  
with Text-to-Speech Given Far-field Acoustic Conditions

○下野晋也, △加藤暁浩, 長野紘之, △後藤悠斗, △能勢将樹(リコー)

- ◆ 本研究では、**遠隔発話音声認識**の End-to-End 学習における **学習データ拡張手法**について論じる。
- ◆ **TTS**により生成された音声を**再生収録**することで、**遠隔発話に変換された合成音声**を獲得する。
- ◆ 前記音声によるデータ拡張が認識精度へ与える影響について、評価結果を報告する。



Fig. 1: A recording equipment for data augmentation.

## 2-2Q-7

### 2-2Q-7 防音用イヤーマフと骨導マイクを用いた 体内音と自己聴取音の模擬

Simulation of body-conducted voice and one's own voice  
by using a soundproofing earmuff and bone conduction microphones

☆陳瑞彦, 西村多寿子, 峯松信明, 齋藤大輔(東大・工)

- ◆ 自己聴取音(one's own voice)を気導音として再現することを試みた先行研究では、気導マイクと骨伝導マイクを用いて気導音と骨導音を同時に収録し、重みをかけて混合する方法が検討されている。
- ◆ しかし、骨伝導による口腔からの音声の伝搬経路は複数存在するため、本研究では、口腔から複数の経路を通して内耳に伝わる振動現象(自己聴取音)から気導音のみを取り除いて定義される振動現象を「体内音」(body-conducted speech)と呼ぶこととする。
- ◆ 「体内音」は空気振動としては観測不可能だが、防音用イヤーマフを装着することで容易に聴取することができる。また、イヤーマフに小型無線イヤホン装着して刺激提示することで、聴取実験における被験者の判断の信頼性を上げることができた。
- ◆ 本研究では、気導音と2種類の骨導音(耳骨導音と喉骨導音)に個別にフィルターをかけ、適切な重みを乗せることで、「体内音」を模擬することを試みた。この場合、音素に依存せず、音源種類と話者にのみ依存する時不変フィルタで、良好な模擬が可能であった。
- ◆ 最後に、「体内音」と気導音を混ぜて自己聴取音を作成し、先行研究の手法で作った自己聴取音との比較実験を行い、本研究の有効性を示すことができた。

### 2-2Q-8

#### 2-2Q-8 フォルマント遷移の話者間正規化による韻の調音特徴の抽出

ベトナム語ハノイ方言を事例として  
☆山岡翔(京都大院・文学研)

◆東アジアから東南アジアにかけて広がる音節基調の孤立語においては、音韻研究の基礎的材料となる離散的表示(いわゆる音素表記や異音レベルの音声表記)のコンセンサスがとれないという、言語研究上の難点を抱えている。

◆その代わりに、当該の言語は音節を基調とするため、音節単位での音声データの収集や分析には長けている。

◆本発表ではベトナム語ハノイ方言を事例として、14名の母語話者の単音節読み上げ音源に含まれる母音フォルマントの遷移全体を話者間正規化することを試みる。

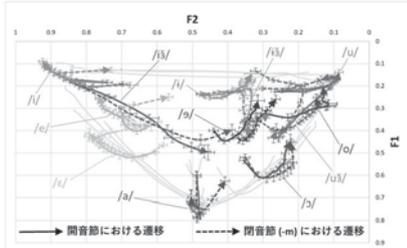


Fig.1: Example of the transition in the vowel space

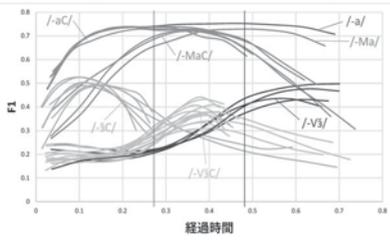


Fig.2: Example of the formant change over time

この正規化遷移データは音韻分析の基礎的材料となるだけでなく、当該言語に本質的な調音特徴について言及しうる点で言語研究上非常に有用であると考えられる。

### 2-2Q-10

#### 2-2Q-10 sinusoidal 合成による波形再現性を有するボコーダに関する検討

A vocoder system for waveform reconstruction using phase information  
☆水谷勇貴(秋田県大), 三浦壮騎(秋田県大院),  
西口正之, 高根昭一, 渡邊貴治, 安倍幸治(秋田県大)

◆近年、声質変換や符号化などの応用を目的とした音声分析/合成の手法として、深層学習を用いるものが主流となりつつある。深層学習を用いることで、高品質な音声合成が可能である一方、音声品質が学習データに依存してしまうことや計算コストが大きくなってしまふなどの問題がある。

◆低演算量かつ高品質な音声合成が可能手法としてボコーダを用いた方法について検討することとした。現在では、位相が音声の品質に及ぼす影響は小さくないということが分かっている。そこで、位相情報を用いることで波形再現性を有するボコーダに関する検討を行った。

◆今回の検討では、位相情報を用いて音声合成を行い、その波形を原音と比較した。その結果、位相情報を用いた音声合成によって、ある程度の波形再現性が得られることが確認された。

◆また、最小位相による合成手法と位相情報を用いた今回の合成手法で合成した音声の聴取実験を行い、聴取者の合成音声についての評価を調査した。

### 2-2Q-9

#### 2-2Q-9 英語学習者と日本語学習者音声に対する多角的評価に関する実験的検討

A study on automatic multi-view assessment of  
L2 English and L2 Japanese speech  
☆榎原卓弥, 藩陽, 斎藤大輔, 峯松信明(東大),  
齊藤一弥(UCL), ボイクマン総子(東大)

◆音声コミュニケーションに焦点を当てて、外国語学習者を評価する場合、様々な評価指標が存在する。例えば、発話の流暢さや訛り度合い、発話内容の理解しやすさ等である。

◆先行研究では、日本人の英語の流暢さを自動推定する技術構築が行われた。しかし、L1=日本語、L2=英語、指標=流暢さ、という限られた条件で行われた実験であった。

◆本研究では、L1/L2/指標を様々な置き換えたコーパス(Tab.1)に対して自動推定を行い、先行研究の汎用性を検討した。

◆また、特徴量として、話速に注目した「発話の量」、母語話者発音との差に注目した「発音の質」の他に、韻律に注目して openSMILE から抽出できる音響特徴量を新たに導入した。

Tab. 1 Corpus

コーパス名	L1	L2	指標
JE1(90/10)	日本語	英語	流暢さ
JE2(125/10)	日本語	英語	流暢さ 訛り度 理解度
PE(90/10)	ポーランド語	英語	流暢さ
MixE(110/10)	様々	英語	訛り度 理解度
MixJ(146/8)	様々	日本語	独白

### 2-2Q-11

#### 2-2Q-11 Attention RBMによる音声特徴量系列の符号化と生成

Representation and generation of acoustic feature sequence of speech based on attention RBM  
◎岸田拓也, 中鹿亘(電通大)

◆系列データの符号化や生成が可能で、少ないデータでも学習可能なモデルを設計することは音声信号処理分野における重要な課題である。

◆再帰的制約ボルツマンマシン (RTRBM) は系列データの特徴を少ないデータで学習できるモデルであるが、系列の長期的な依存関係をとらえることは困難である。

◆そこで、自己注意機構を付与することで、長期的な依存関係を捉えることができるRBM, Attention RBMを提案した。

◆Attention RBMは音声特徴量系列に対して、RTRBMよりも安定して学習ができ、特徴量の符号化・再構成でもRTRBMを上回る性能が得られた。

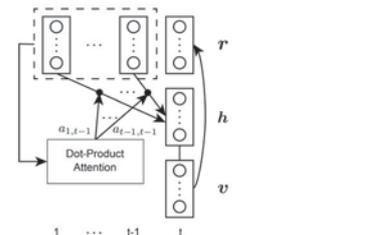


Figure: Graphical representation of the proposed model.

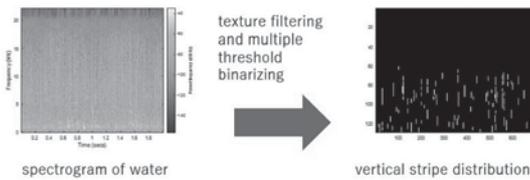
### 2-4P-1

#### 2-4P-1 テクスチャー認識をスペクトログラムに適用した水流音等の連続的な環境音の認識

Continuous environmental sound recognition based on spectrogram texture recognition

○金澤坤, △佐々木健

- ◆ Spectrogram of continuous environmental sounds such as flowing water sound has many vertical stripes.
- ◆ Texture filtering and other image processing methods were applied to spectrogram images to extract vertical stripe distribution.
- ◆ Several statistics of vertical stripe distribution were calculated as sound features of sounds such as flowing water, rustling sound of trees, engine, and crackling sounds in a bonfire.



- ◆ Statistical features were; distribution of average time interval between stripes, average number of stripes in the same time frame, average stripe length, frequency centroid of stripes, occupancy rate of stripes in the spectrogram. Statistical features were compared and discussed. The proposed feature values were different for all sound categories, especially for sound of flowing water.

### 2-4P-3

#### 2-4P-3 「ゾクゾク」感を伴う音に関する研究

An investigation on sound with a "zoku-zoku" sensation.

○寺岡諒, △向中野公太, △寺本涉 (熊本大)

- ◆ ASMR (ささやき声等の音によって生じるゾクゾクする感覚) は、主観的なリラククス効果や幸福感上昇に加えて、心拍数の低下等の生理反応を引き起こす [Barratt & Davis, 2015; Poerio et al., 2018]
  - ASMR を誘発する音が心拍数に及ぼす影響の経時変化を検討した
- ◆ 3分間の運動で聴取者の心拍数を上昇させた状態で、ASMR が生じやすい音 (高 ASMR 音)、生じにくい音 (低 ASMR 音) をそれぞれ聴く条件、無音条件で心拍数の経時変化を比較した
- ◆ 実験の結果、高 ASMR 音を聴いた場合、他の条件に比べて有意に早く心拍数が減少した
- ◆ しかし、運動前と音聴取後に気分を評価させたところ、高 ASMR 音でのみ快適性が低下し、高 ASMR 音以外で覚醒度が低下した
  - ASMR による心拍数の減少は、音による気分の改善 (快適性の向上) とは関係なく生じることを示唆

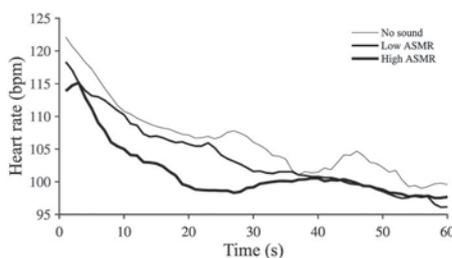


図1: 各音条件における経過時間と心拍数の関係性 (N = 19).

### 2-4P-2

#### 2-4P-2 拍分割にともなうテンポ変化知覚: 複数テンポでの比較

Tempo perception of stimuli with beat divisions : Comparison of multiple tempi

☆山岡卓哉, 饗庭絵里子(電通大), 橋亮輔(東大)

- ◆ 音楽演奏に重要な能力のひとつとしてテンポキープがあげられる。特に、複数人で演奏する場合には、自身のパートのリズムパターンが変化しても同じテンポを保たなければならない。テンポをキープするためには、拍の時間長を正確に知覚し記憶する必要があるが、物理的な時間長と知覚される時間長は必ずしも一致しない。また、実験に用いられるテンポの平均 (グローバルテンポ) に依存して、知覚に変化が生じることも知られている。
- ◆ そこで本研究においては、実際の音楽と同様に途切れなく一定間隔で音が提示されてテンポが知覚されるような刺激で、かつ単純な拍分割があるリズムパターンを用いて、グローバルテンポがテンポ変化の知覚にどのような影響を及ぼすのかを検証した。

- ◆ その結果、相対的に基準テンポが速い場合には速いと知覚し、基準テンポが遅い場合には遅いと知覚する傾向にあることが示された。また、拍分割した音刺激のテンポ知覚は、直前に聞いたテンポに影響されることが示唆された。

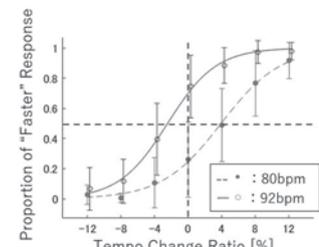


Fig. 1 Experimental Results

### 2-4P-4

#### 2-4P-4 交通音の距離感の知覚における老人性難聴の影響: 模擬難聴システムによる検討

Effects of age-related hearing loss on perception of the distance to approaching car : Investigation using hearing impairment simulator.

◎古屋孝基, 渡邊優也, 松井淑恵 (豊橋技科大院)

- ◆ バイノーラル録音した交通音の聴取実験により、老人性難聴が交通音の距離感の知覚に与える影響について模擬難聴を用いて検討した。
- ◆ 交通音の加工条件として、録音時の音圧を再現した音 (NH 条件)、70 代男性の聴力レベルを圧縮特性健全度 50%、100% で模擬した音 (HIS50% 条件、HIS100% 条件)、HIS50% 条件および HIS100% 条件と平均音圧レベルが同じになるよう音圧を低減させた音 (FA50% 条件、FA100% 条件) の 5 種類を設定した。
- ◆ 実験参加者は交通音を聴取した後、仮想空間上に表示された車両の位置を調整することで知覚された車両までの距離を回答した。
- ◆ 結果、老人性難聴により、車両までの距離が健聴者よりも遠くに知覚され、圧縮特性の変化も距離の知覚に影響することが示唆された。

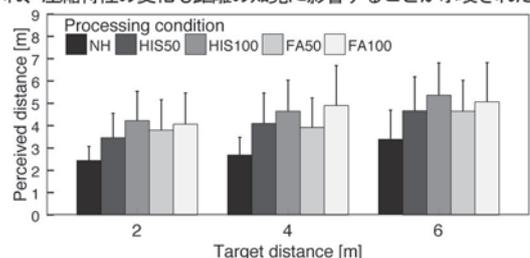


Fig.1: Perceived distance to the approaching car (10-20 km/h). The horizontal axis is distance between subjects and centerline of car in real space. The vertical axis is perceived distance in virtual space. Error bars represent standard error. The colors represent processing condition.

### 2-4P-5

#### 2-4P-5 駅構内の残響・雑音下での緊急放送に対する心理的評価: 模擬難聴システムによる老人性難聴の影響の検討

Psychological evaluation of speech announcement for emergency under reverberation and noise inside a station: effect of age-related hearing loss using a hearing impairment simulator.

☆渡邊優也、古屋孝基(豊橋技科大院)、

高橋慧、中家諒(TOA)、松井淑恵(豊橋技科大院)

- ◆高齢者実験の代わりに模擬難聴システムを利用して、公共空間における緊急放送の聞こえがどのような要因(残響、雑音、放送音声そのもの)に影響されているか、また、そこに聴取者の聴力が加わったときに、どのような変化が起きるかを調査することを目的として実験を実施した。
- ◆公共空間は鉄道駅構内で、発話訓練を受けていない人によって緊急放送が行われることを想定した。
- ◆著者の先行研究で音声収録・評価実験、各要因を検証する加工音声評価実験を実施した。本稿では条件を絞り込んで残響・SNR・模擬難聴条件を組み合わせて評価した。
- ◆聴取実験の結果、残響を抑えることによって、聞き取りやすさを向上させることができるが、SNR が低いと残響の有無に関わらず聞き取りやすさが低下すること、老人性難聴者は、残響やSNRによる放送音声の聞き取りやすさ・緊急性的変化を健聴者ほど感じにくくなること、SNRの上昇による緊急性的向上は、残響の種類に依存すること、残響による緊急性的印象向上の効果は、残響の長い音声の場合、老人性難聴者には現れない可能性などが示唆された。

### 2-4P-7

#### 2-4P-7 同一話者による怒り音声と喜び音声のモーフィングに対する感情知覚

Perception of emotion for morphing speech between angry and joy by the same speaker

☆坂本拓也、坂下尚史(豊橋技科大)、

河原英紀(和歌山大)、松井淑恵(豊橋技科大)

- ◆「平静」音声と「喜び」音声をモーフィングしたときに、モーフィング割合に応じて聴取者がどの程度喜びの感情を知覚するのかを測定し、発話者の性別による違いや個人差を検討した。
- ◆「怒り」と「喜び」の音声のモーフィングによってどの感情が知覚されるか、知覚される感情間にもどのように影響し合うかを調査した。
- ◆「平静」「怒り」「喜び」の感情で発話された男声と女声それぞれ1名の「はい」を原音声として、STRAIGHT モーフィングを利用し、「平静」から「喜び」、「喜び」から「怒り」の中間の音声を作成した。喜び音声の割合を0、20、40、50、60、80、100[%]に設定した。
- ◆「平静」から「喜び」の音声では、50%ずつの割合でモーフィングした音声とそれ以外の割合の2つのモーフィング音声を聞き、どちらがより喜びの感情に聞こえるか GUI 上で実験参加者に判定させた。喜びの割合に応じて喜びが知覚される割合は単調増加したが、その増加具合には個人差が存在することが示唆された。
- ◆「喜び」から「怒り」の音声では、どの感情に聞こえるかを複数の感情の選択肢から回答させた。ほとんどの参加者で喜びと怒りが多く知覚されたため、上記と同じ実験を「喜び」と「怒り」の両方の感情を対象として実施した。平静が怒りに与える影響と喜びが怒りに与える影響の違いがあるが、その差は大きくないことがわかった。

### 2-4P-6

#### 2-4P-6 氣息性の制御による擬似 breathy voice の生成と聴取評価

Synthesis of breathy voice by manipulating parameters related to breathing and its subjective evaluation

☆葉玉大智、松井淑恵(豊橋技科大院)

- ◆通常の発声(modal voice)から無声の発声(whisper voice)までシームレスな声質変換を可能とするため、modal と whisper の中間の声質といえる breathy voice の生成法を提案し、生成法の変換補正量を変化させることで、modal から whisper まで段階的な声質変化を試みた。
- ◆収録したbreathy voice と modal voice から抽出した平均スペクトル包絡と平均非周期性指標の差分をもとに擬似 breathy voice を生成した。
- ◆擬似 breathy voice を用いて、modal から whisper まで段階的に声質変化させた際の印象変化を形容詞対によって評価したところ、擬似 breathy voice は実際の breathy voice よりもかすれており、やわらかさが足りないが、声の息っぽさは徐々に増していることが確認された。

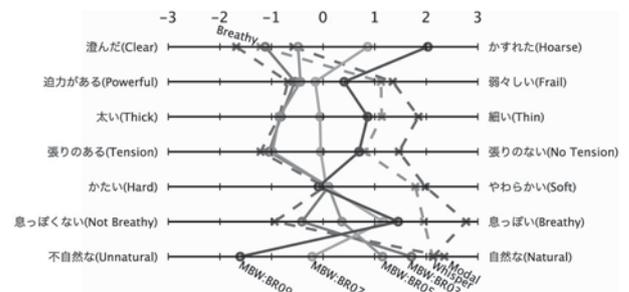


Fig.1: Changes in the impression of MBW (Modal to Breathy to Whisper) speech sound by BR (Breathing Rate). (Participants ×: Original, ○: MBW. BR 0: modal ←→ 10: whisper)

### 2-4P-8

#### 2-4P-8 精神疲労が聴性誘発反応とオリーブ蝸牛束反射に与える影響の比較

Comparison of Effects of Mental Fatigue on Auditory Evoked Response and Medial Olivocochlear Reflex

☆鈴木裕登、大塚翔、中川誠司(千葉大)

- ◆脳幹から外有毛細胞へは、オリーブ蝸牛束と呼ばれる遠心性神経の投射がある。オリーブ蝸牛束は音に反応し、外有毛細胞の増幅を抑制することで内耳を保護している。そのため、この反応(オリーブ蝸牛束反射: MOCR)の強度を測定することで、騒音性難聴のリスクを事前に評価できる可能性がある。
- ◆しかし、MOCR 強度は測定度に変動するため、騒音性難聴の正確なリスク評価は困難である。MOCR 強度が変動する要因に関連して、著者らはこれまでに、精神疲労時に MOCR が増強することを報告してきた。
- ◆本研究では、MOCR が精神疲労によって変動するメカニズムを解明するため、聴性誘発反応を計測することで脳幹から皮質にかけての脳活動を評価し、MOCR との関係調べた。
- ◆精神疲労は、フリッカー検査(CFF)によって客観的に評価し、MOCR は耳音響放射(外有毛細胞の増幅を反映)を測定することで評価した。
- ◆計算課題の繰り返しに伴い、臨界フリッカー周波数(CFF)は減少した。計算課題によって精神疲労が増加したことを示している。
- ◆計算課題によって、MOCR と聴性誘発頭頂部緩反応(SVR:皮質活動を反映)が増加する傾向が見られた(Fig.1)。この結果は、精神疲労に伴う MOCR の増加に、皮質活動の増加が関与している可能性があることを示唆している。しかし、皮質活動と MOCR で、増加が見られる試行数が異なったため、今後は両者の関係を精査していく必要がある。

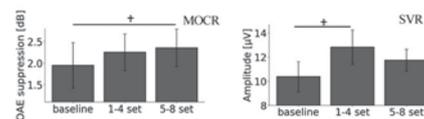


Fig.1 Calculation task-induced changes of MOCR and SVR relative to its baseline. † p<0.1.

### 2-4P-9

#### 2-4P-9 規則性に基づく予測によってオリブ蝸牛束反射が変動するメカニズムの解明 -脳波とオリブ蝸牛束反射の比較-

Mechanisms Underlying the Enhancement of Medial Olivocochlear Reflex Associated with Rhythm-based Prediction  
: An investigation on Correlative changes of Electroencephalogram and Medial Olivocochlear Reflex

○石坂勇毅, 大塚翔, 中川誠司(千葉大)

- ◆ 脳幹から外有毛細胞へは、オリブ蝸牛束と呼ばれる遠心性の神経投射が存在し、音響刺激に反応して、外有毛細胞の増幅を抑制し、雑音下での聞き取りを改善する役割や内耳を強大音から保護する機能を果たしていると考えられている。この反応はオリブ蝸牛束反射 (MOCR) と呼ばれている。
- ◆ 著者はこれまでに、著者はこれまでに、先行する音列が規則的に刺激音出現のタイミングが予測できる場合に、MOCR が強くなることを報告してきた。本研究では、MOCR と皮質の反応を同時計測することで、予測によって MOCR 強度が増加するメカニズムを検討した。
- ◆ 不規則なタイミングで到来する刺激の中に (Irregular 条件)、規則的なタイミングで到来する刺激を 40% の割合で呈示し (Regular 条件)、MOCR、および、N1-P2 の振幅値の時間変動を比較した。
- ◆ Irregular 条件と Regular 条件での MOCR 強度の差 ( $\Delta MOCR_{reg}$ ) と CAEP の N1-P2 の振幅値の時間変動には負の相関があることが示された (Fig.1)。この結果は、予測に伴う MOCR 強度の増加に聴覚皮質の活動が関与していることを示唆している。

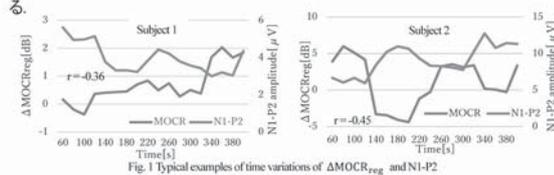


Fig. 1 Typical examples of time variations of  $\Delta MOCR_{reg}$  and N1-P2

### 2-4P-11

#### 2-4P-11 注意の時間的変動がオリブ蝸牛束反射に与える影響

Effect of temporal fluctuations in attention on medial olivocochlear reflex

○鈴木麻里子, 石坂勇毅, 大塚翔, 中川誠司(千葉大)

- ◆ 脳幹から外有毛細胞へは、オリブ蝸牛束と呼ばれる遠心性神経の投射がある。オリブ蝸牛束は音響刺激に反応し、外有毛細胞の増幅を抑制する働きがある。この抑制反応は、オリブ蝸牛束反射 (MOCR) と呼ばれ、内耳を保護する役割を果たしている。MOCR 強度を測定することで、騒音性難聴のなりやすさを事前評価する手法への応用が期待されている。
- ◆ しかし、MOCR 強度は測定のために変動するため、安定した評価が困難であり、そのメカニズムも未解明であった。これまでに、MOCR が聴覚的・聴覚的注意の影響を受けることが報告されているものの、注意・非注意の条件間での比較に留まっており、注意状態を定量的に評価した研究はなかった。
- ◆ 本研究では、注意の変動を、聴覚課題に対する反応時間を測定することで定量的に評価し、MOCR との関係調べた。MOCR は耳音響放射 (外有毛細胞の増幅を反映) を測定することで評価した。
- ◆ 聴覚課題に対する反応時間が小さいほど、MOCR 強度が増加した (Fig. 1)。反応時間が小さいほど、注意レベルが高いと解釈されることから、この結果は、聴覚に対する注意レベルが高いほど、MOCR 強度が強くなることを示唆している。

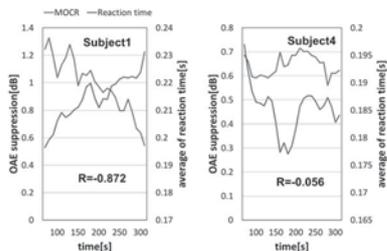


Fig. 1 Typical examples of time course of changes in reaction time and medial olivocochlear reflex strength

### 2-4P-10

#### 2-4P-10 Deep Prior を用いた聴覚誘発反応の電流源推定

Current source estimation of auditory evoked response using deep prior

☆山名莉央, 矢野肇, 高島遼一, 滝口哲也(神戸大), 中川誠司(千葉大)

- ◆ 脳機能計測手法である脳磁図、脳波の計測結果から脳内の電流分布の推定は劣決定問題であり、解くためには電流源の分布を明示的に与える必要があるがその分布は不明である。
- ◆ 本研究では明示的に事前分布を与えないようにするため、Deep Prior を用いて電流源推定を行った。
- ◆ Deep Prior は畳み込みからなるネットワークが事前分布の役割を持つという考え方に基づいた手法であり、様々な逆問題を解くために応用されている。
- ◆ 畳み込み構造を持つネットワークが電流源の事前分布を表現可能なことを、シミュレーションと実際の聴覚誘発反応の MEG データを用いて示した。

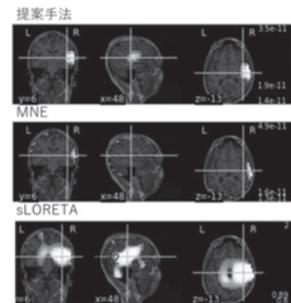


Fig. 1 Current source densities estimated from an auditory evoked field

### 2-5P-1

講演取消

### 2-5P-2

#### 2-5P-2 仮想音環境における残響空間下での先行音効果の検討

Study of evaluation characteristics of preceding effect using virtual sound environment

☆武藤泰斗, 大隅歩, 伊藤洋一(日大・理工)

- ◆地下街や大規模建物内で避難誘導を積極的に行うための「先行音効果を利用した音声避難誘導システム」の研究を行っている。このシステムを実装する場合、基本音響設計に加えて現場での調整が必要であり、これをVirtual Reality技術で支援することを考えている。
- ◆本報告では、残響時間の異なる実空間を仮想空間上に再現し、これら空間における先行音と後続音の融合と分離、及び音の方向感強度について比較検討した。

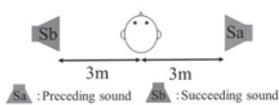


Fig. 1 Layout of experiment

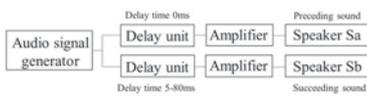


Fig. 2 Block diagram

Table 1 Evaluation index

- 評価値7: Sa側に非常に方向感がある
- 評価値5: Sa側に方向感がある
- 評価値3: Sa側に方向感が少しある
- 評価値1: 方向感が全くない分難してわからない

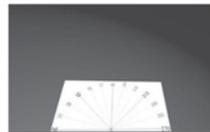


Fig. 3 Visual conditions

### 2-5P-4

#### 2-5P-4 移動ロボットを用いた多点室内音響測定の評価

An evaluation on room acoustic measurement at multiple points by using mobile robot

☆上原正志, 石川直生(千葉工大), 福島学(日本文理大), 大川茂樹(千葉工大)

- ◆ 移動ロボットを用いた多点での室内音響測定
  - 測定の効率化, 簡易化
  - 本稿では位置検出精度についての検討
- ◆ 運用の仕方によって位置検出精度に影響
  - 特定の環境下では不偏標準偏差±0.025[m]

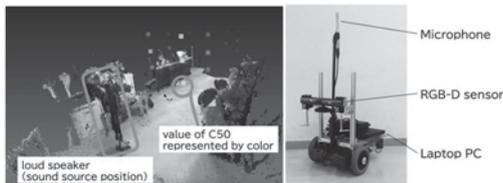


Fig. 1 room acoustic measurement by using a mobile robot (left: Visualization example: distribution of C50, right: overview of the mobile robot)

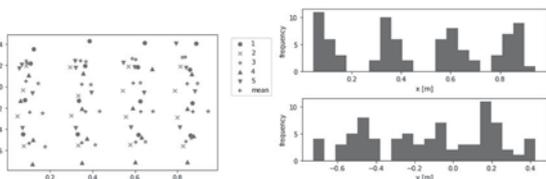


Fig. 2 results of evaluation on estimating measurement positions (left: distribution of estimated measurement positions, right: histogram of estimated measurement positions)

### 2-5P-3

#### 2-5P-3 再生核を用いた瞬時到来パワー予測手法の検討

Instantaneous Arrival Power Prediction Method Using Reproducing Kernel

◎岩見貴弘, 尾本章(九大芸工)

- ◆一般次元に於いて、超球に制限した帯域制限空間を構築し、その再生核を導出した。
- ◆この再生核の線形和で音場を記述することにより、任意配置のマイクロアレイを用いて瞬時到来パワーを推定し、到来方向を推定する手法を検討した。
- ◆数値実験より、空間エイリアシングの生じない周波数に於いて、周波数が高いほど良い到来方向推定結果を示した。
- ◆適切な波数依存重み関数を用いる事で推定性能が向上する可能性を示した。
- ◆提案手法は瞬時情報のみから推定を行うため、前後の違いを捉える事が出来ないという欠点があるが、これは今後の課題としたい。

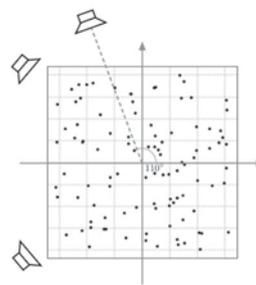


Fig. 1 Sensor placement

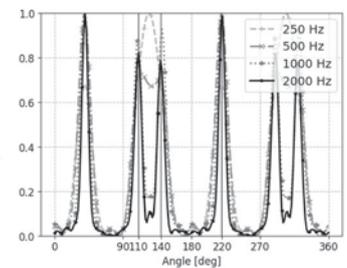


Fig. 2 Experimental results of DOA estimation using proposed method

### 2-5P-5

#### 2-5P-5 ロボットによる室内音響測定のための経路決定

E A Study of Routing Algorithm for Room Acoustic Measurement by Robot

☆石川直生, 上原正志, 大川茂樹(千葉工大), 福島学(日本文理大)

- ◆移動ロボットを室内音響測定の受音装置移動に利用
- ◆一度に計測可能な範囲を変化させた際の測定値の誤差, 所要測定点数についてシミュレーション環境にて調査
- ◆2種類の条件で行った結果, 明らかな差が見られた

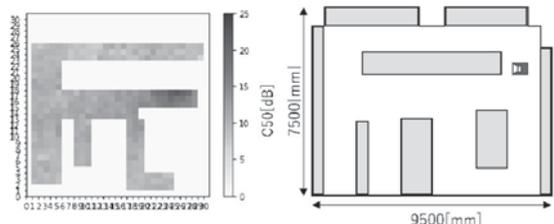


Fig. 1 Measurement result of C50 in real environment (left), Floor plan of measurement environment (right)

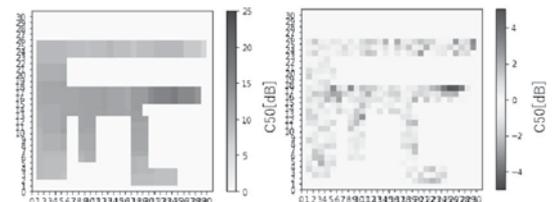


Fig. 2 Measurement result of C50 in simulation (left), Map of measurement error about C50 (right)

## 2-7P-1

### 2-7P-1 予測歪み信号加算による パラメトリックスピーカの歪み低減の検討 —距離依存性の評価—

Reduction of harmonic distortion in parametric loudspeaker by addition of predicted distortion signal using: Evaluation of distance dependence

☆原田俊平, 野村英之(電通大)

- ◆パラメトリックスピーカの音質劣化の原因の一つとして、高調波歪みが指摘されている。
- ◆入力信号の高調波に相当する信号を加算する高調波歪み低減法をシミュレーションで評価し、距離依存性について検討した。
- ◆単側波帯振幅変調を採用し、搬送波周波数を 40 kHz、入力信号周波数を 1 kHz とし、歪み率は第 2 高調波 (Second Harmonic Ratio, SHR)

$$SHR = \frac{P_{2\text{ kHz}}}{P_{1\text{ kHz}}} \times 100 [\%]$$

で評価した ( $P_{i\text{ kHz}}$  は  $i\text{ kHz}$  の音圧である)。

- ◆Fig. 1 にその結果を示す。提案法により遠距離場での歪みが低減できること、音源面音圧が高いと、より近距離から低減が行えることが確認できた。

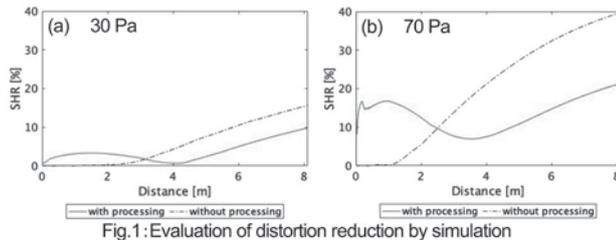


Fig.1: Evaluation of distortion reduction by simulation

## 2-7P-3

### 2-7P-3 気泡の初期成長過程における ソノルミネッセンスの周波数依存性

Frequency dependence of sonoluminescence during the initial growth process of bubbles

☆中村悠人, 乾綾華, 杉野史弥, 山本健(関西大)

- ◆従来の初期成長過程の MBSL の観測はパルス(バースト)超音波を使用したものであり、超音波を連続的に照射した研究報告は少ない。
- ◆本研究の目的は、超音波周波数の違いが MBSL に与える影響を初期気泡成長の観点から把握することである。
- ◆2 周波(410 kHz, 1.0 MHz)の連続的な超音波を照射し、気泡の初期成長過程の MBSL 強度の時間変化を測定した。
- ◆MBSL 強度が定常状態に至るまでの時間を比較した。
- ◆周波数 1.0 MHz の方が 410 kHz と比較して、MBSL 強度が低く、発光するまでの時間に遅延が確認された。

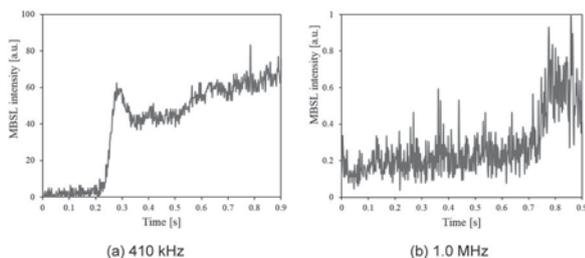


Fig. 1 Time variation of MBSL intensity at frequencies of (a)410 kHz and (b)1.0 MHz.

## 2-7P-2

### 2-7P-2 造影剤の破壊・凝集現象がドプラ効果に 及ぼす影響の基礎検討

Fundamental study of the effect of the destruction / aggregation phenomenon of microbubble on the Doppler effect

☆木戸口哉太(千葉大・工), 大村真朗(富山大・工),  
山口匡(千葉大・CFME), 吉田憲司(千葉大・CFME)

- ◆リンパ管を可視化する手法として、音響放射力によって生じる造影剤の移動をドプラ法により定量する動的造影超音波法を提案している。
- ◆高音圧条件で生じる造影剤の破壊現象や凝集現象が起こると、移動速度の誤推定や実用上で繰り返し検査が困難になるなどの問題が生じる。
- ◆造影エコーも信号強度は造影剤濃度に比例することから特に信号強度に着目し、造影剤の破壊が生じる音圧閾値について検討した。
- ◆超音波パルスを繰り返し周波数 5 kHz で 1000 回照射し、1 波目と 2 波目の受信エコーの信号強度の減少率を評価した。
- ◆Mechanical Index (MI) 0.07 - 0.61 の範囲で同様に評価を行い、信号強度の減少率の音圧依存性を評価した (Fig. 1)。
- ◆信号強度の減少率が急激に増加する MI 0.3 程度が音圧閾値と考えられる。

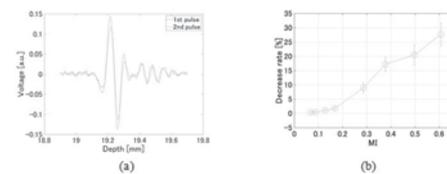


Fig.1: (a) An example of received echo signal at 1st pulse and 2nd pulse and (b) decrease rate of echo intensity.

## 2-7P-4

### 2-7P-4 巨大ベシクルに内包されたトレーサー粒子 の放出制御に関する基礎検討

Fundamental study on controlled release of tracer particles contained in giant cluster vesicles

☆瀬尾康太(千葉大・工), △章逸汀(千葉大・院融合),  
△豊田太郎(東大・院総合文化), △林秀樹(千葉大・CFME),  
山口匡, 吉田憲司(千葉大・CFME)

- ◆病巣マーカーおよびセンチネルリンパ節 (SLN) トレーサーとして機能する多機能ベシクル凝集体を開発している。
- ◆超音波照射により GCV を破壊し、SLN トレーサーとして内包される蛍光リポソームを放出する手法を提案している。
- ◆超音波照射時の GCV の挙動を光学的に観察し、超音波照射後に回収した試料液を対象に、吸光スペクトルを計測した[Fig.1]。
- ◆超音波造影剤存在下 (UCA(+)) では、超音波照射時に GCV の破壊が促されていることが観察像から確認でき、また蛍光リポソームの吸光波長において吸光度が高値となった。
- ◆超音波造影剤が存在しない場合 (UCA(-)) では、GCV 挙動には変化がなく、吸光度は低値であった。
- ◆超音波造影剤と超音波の相互作用を利用することで、GCV の破壊と内包物質の放出を促進できることが示された。

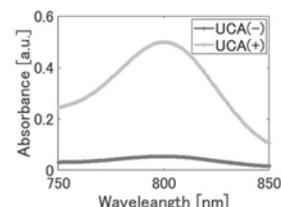


Fig.1 Absorbance spectrum in cases of UCA(+) and UCA(-)

## 2-7P-5

### 2-7P-5 超音波による *E. coli* 及び *S. cerevisiae* の選択的不活性化

Selective inactivation of *E. coli* and *S. cerevisiae* by ultrasound

☆西口慶, 上畑峻也, 橋本柚子, 山本健(関西大)

- ◆超音波による微生物の不活性化は微生物単体を対象としたものが多く, 不特定多数の菌が混在する一般雑菌を除いては, 複数の微生物を混合したサンプルに対する研究は多くはない。
- ◆本研究では, *E. coli* 及び *S. cerevisiae* を混合した懸濁液への超音波照射における選択的な不活性化の可能性について検討した。
- ◆周波数 430 kHz 及び 1600 kHz において *E. coli* 及び *S. cerevisiae* の選択的不活性化に成功した。430 kHz では *E. coli* だけを選択的に不活性化することが可能であり, 1600 kHz では, *S. cerevisiae* を優先的に不活性化させることができた。(Fig. 1)

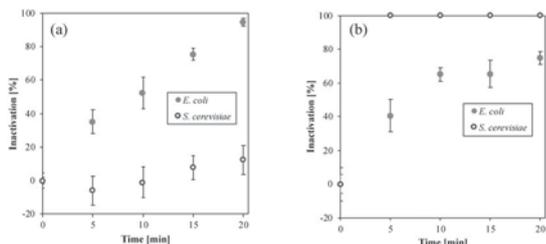


Fig. 1 Inactivation of *E. coli* and *S. cerevisiae* in mixed cell suspension against irradiation time at 430 (a) and 1600 (b) kHz.

## 2-7P-7

### 2-7P-7 音響放射力によって加振された膜の振動特性の測定と評価

Measurement and evaluation of membrane vibration

excited by acoustic radiation force

☆北村香子, 野村英之(電通大)

Vibro-acoustography において, 画像化に用いる信号は, 対象物の振動から生じる音響放射を周囲の媒質を介して測定しているため, 生成された画像が対象物の振動特性を反映しているかについて検討する必要がある。

本報告では, その検討を行う前段階として, 周端を円形状に固定したラップの振動波形をレーザー Doppler 振動計によって取得することで, その周波数特性を測定した。また, その周波数特性を対象物の固有振動数の理論値と比較した。

Fig. 1 に振動の周波数特性を示す。周波数特性のピークと固有振動数の理論値は, どちらも高周波ほど, それらが現れる周波数間隔に広がりが見られ, 傾向として一致していた。得られた周波数特性は確かに対象物の振動特性を示していることが分かった。

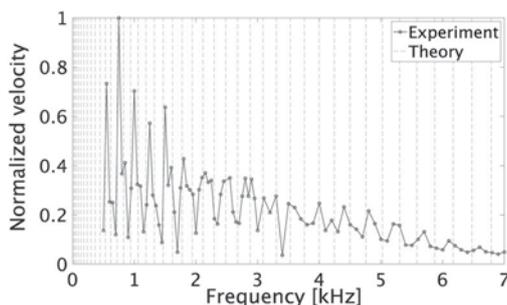


Fig.1: Frequency characteristic of membrane vibration excited by acoustic radiation force

## 2-7P-6

### 2-7P-6 40 kHz 超音波の気導聴覚閾値の推定

Hearing threshold for 40 kHz airborne ultrasounds

☆土井 公一朗, 石川 大夢(千葉大),

長谷 芳樹, 星 貴之(ビクシーダストテクノロジーズ(株)), 中川 誠司(千葉大)

- ◆空中超音波フェーズドアレイによって, 頭部に 40 kHz の強力超音波を照射すると, 甲高いピッチが知覚されることがある。
- ◆空中超音波知覚の基礎検討として, 超音波アレイスピーカを用いて 40 kHz 空中超音波の検出閾を計測した。
- ◆全ての被験者が刺激音を検出可能であった。内耳がこの帯域の超音波に対する感度をある程度有していることが改めて示された。
- ◆一方, 平均検出閾は 133.3 dB SPL であり (Fig. 1), 先行研究で示された 24 kHz 超音波の検出閾 (平均 88 dB SPL) と比較しても, 大きな値となった。超音波周波数の増大に伴う内耳の反応特性の低下, もしくは中耳の伝達特性の低下が示唆される。

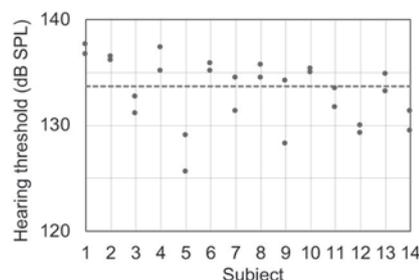


Fig.1: Hearing threshold for 40-kHz airborne ultrasound calibrated by B&K4939 with protection grid. The dotted line represents the average among subjects.

## 2-7P-8

### 2-7P-8 水熱合成法により製膜した圧電結晶膜の特徴とその応用

Evaluation of high intensity and high frequency ultrasonic transducers using KNbO<sub>3</sub> piezoelectric films

○石河睦生, 齋藤 直(桐蔭横大), 田原麻梨江, 黒澤 実(東工大)

- ◆水熱合成法により製膜した KNbO<sub>3</sub> 圧電結晶膜を用いて高周波高出力動作が可能な超音波トランスデューサを試作した。
- ◆水中において試作した超音波トランスデューサの送信音圧の測定では音圧 4MPa の送信をハイドロフォンを用いて確認した。また高音圧超音波の送信時, 大きな波形ひずみと偶数倍奇数倍の高調波が検出され, 音響流の発生も確認された。
- ◆KI 溶液を用いてキャビテーション発生の有無について測定した。照射時間の増加と共に 355nm の吸光度が増加されていることが分かった。試作した超音波トランスデューサによる強力超音波照射により, I<sup>-</sup>イオンが KI 溶液中に増加したことが考えられる。
- ◆10MHz の高周波キャビテーションによるソノルミネッセンスの確認を行った。観察により得られた写真を Fig. 1 に示す。開口部近傍に発光が観察された。
- ◆高周波キャビテーションによる対象物の洗浄効果の評価を行った。超音波洗浄機よりも 1/10 程度の出力であるにもかかわらず洗浄効果が期待できることが分かった。

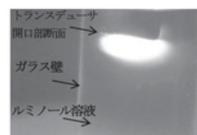


Fig. 1: 観察されたソノルミネッセンス

2-7P-9

2-7P-9 リチウムナイオベート素子を用いた超音波噴霧に関する基礎的検討

Basic study on ultrasonic atomization using lithium niobate

☆吉岡圭佑(東京理科大), △倉科佑太(東工大),  
△小川愛実(慶応大), 朝倉巧(東京理科大)

- ◆超音波噴霧における素子としては PZT が代表的であり、広く一般的に使用されている。しかし、その代わりにリチウムナイオベート(LN)を素子として用いることで、効率的に噴霧が可能であることが明らかになったが、その噴霧特性については多くの知見が得られていないのが現状である。
- ◆本検討では、大きさの異なる素子を5種類用意し、素子の大きさの違いによる噴霧特性について検討した。1つ目の検討として、噴霧の様子をハイスピードカメラで撮影し、ミストの粒径について解析した。
- ◆2つ目の検討として、連続的に噴霧できる機構を作成し、それぞれの素子で1分あたりの噴霧量を計測した。さらに素子面積と消費電力の面から噴霧効率について評価した。

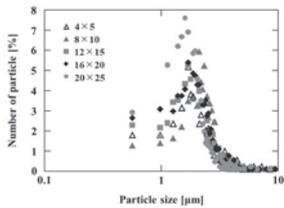


Fig. 1 Particle size distribution of ultrasonic mist

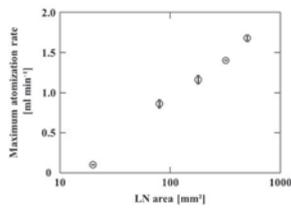


Fig. 2 Relationship between LN area and maximum atomization rate (n = 5, Mean ± S.D.)

2-7P-11

2-7P-11 バルク厚み振動子による液中推進システム

Underwater propulsion system via thickness-vibration-mode transducer

◎孔徳卿, 花田徳(室蘭工大院), △銭源, 黒澤実(東工大), 青柳学(室蘭工大院)

- ◆本研究の目的は、固体から液体へと波動が放射するときに、界面に生ずる音響放射圧を音響放射推進力として利用し、ヒレやスクリューといった可動部を持たない、液中アクチュエータの開発である。厚み振動子の片側に空洞が設計され、水中で一方音響放射推進力を生み出す推進システムを検討する。
- ◆Fig. 1 に示す厚み振動子の片側に空洞が設計され、水中で一方音響放射推進力を生み出すスイマーを設計した。円板状 PZT 厚み振動子の厚み振動モードを用いる。共振周波数は 2.065MHz, 円板の直径は9mm, 厚さは1mm である。3D プリンターで作製したアクリルプロトタイプスイマーの寸法は18×14mm である。
- ◆Fig. 2 に示すように駆動電圧が12.4 V<sub>pp</sub> のとき、水中で0.2 mN の0速度推進力が測定された。Fig. 3 に示すように6.1 mm/s の速度が確認された。したがって、推力が0.2 mN 以上で、スイマー推進システムが実現された。



Fig. 1 Prototype swimmer.

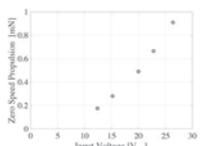


Fig. 2 Zero speed propulsion measurement

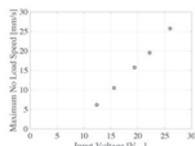


Fig. 3 No load speed.

2-7P-10

2-7P-10 水熱合成法を用いた KNbO<sub>3</sub> 圧電薄膜のアドミタンス特性評価

Evaluation of piezoelectric characteristics of KNbO<sub>3</sub> thin film by hydrothermal Method

☆村瀬貴一, △藤井健人, 田原麻梨江(東工大)  
△齋藤直, 石河睦生(桐蔭大)

- ◆水熱合成法による KNbO<sub>3</sub> 圧電薄膜は高周波かつ高出力で駆動可能な圧電材料であり、治療用超音波トランスデューサやバイオプリンティングへの応用が期待できる。
- ◆KNbO<sub>3</sub> 圧電薄膜と電極とからなる二種類の簡易構造の超音波トランスデューサを新規に製作し、従来困難であった KNbO<sub>3</sub> 圧電薄膜単体としての圧電特性を測定し、等価回路を用いた計算結果と比較した。
- ◆SMA コネクタ付きのトランスデューサは、コネクタ等の電気的な共振が大きいため、圧電体由来の共振反共振のピークが不明瞭であったが、コネクタなしのトランスデューサは KNbO<sub>3</sub> 単体薄膜の共振反共振がより鮮明に確認でき、共振周波数は計算値とおおよそ近い値を示した。

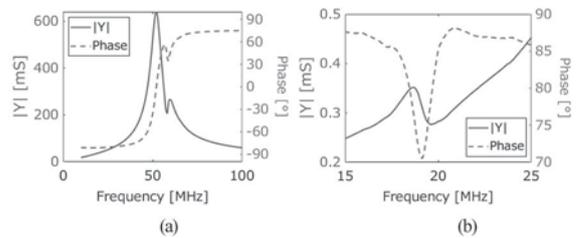


Fig. 1: Measured admittance characteristics; (a) with connector and (b) without connector

2-7P-12

心拍による頸動脈変位の2次元超音波計測におけるスライス方向変位の影響に関する検討

Investigation on effects of elevational motion on 2D ultrasonic measurement of displacement of carotid arterial wall

☆長尾生成, 石川数馬, 茂澄倫也, 大村眞朗,  
長岡 亮, 長谷川英之(富山大工)

- ◆2次元位相差追跡法は、高速超音波イメージングにおいて、高い計算速度と高い精度で2次元(ラテラル方向x, 距離方向z)の変位・速度推定を行う際に有用である。本報告では、2次元変位・速度推定におけるスライス方向変位の影響を検討するために、スライス方向への平行移動を与えたファントムを作成し、計測の精度を比較する。
- ◆頸動脈を模した円筒状のファントムに、中心からの拍動による血管壁の速度を与え、4つの異なるスライス方向への平行移動速度を与え、2次元位相差追跡法により推定した速度の誤差を比較した。
- ◆Fig. 1 にファントムの B モード像と速度推定をした関心点の範囲の図、Fig. 2 に推定誤差のグラフを示す。Fig. 2 より、スライス方向変位が大きくなるほど、推定誤差が大きくなる事が分かる。

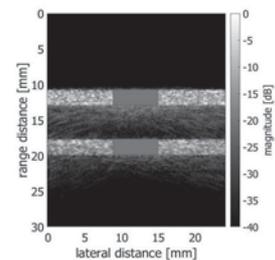


Fig. 1: シミュレーションファントムのBモード像と関心点.

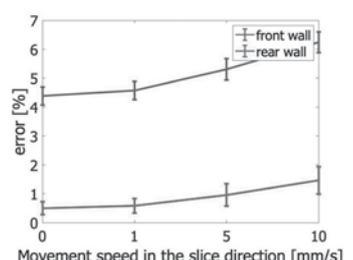


Fig. 2: 各スライス方向変位速度における距離方向速度の推定誤差.

### 2-7P-13

#### 2-7P-13 集束超音波ビームを用いた開口合成法による血流測定法の検討

Investigation on blood flow measurement by synthetic aperture method with focused ultrasound beam

☆杉岡虎太郎, 茂澄倫也, 大村真朗, 長岡 亮, 長谷川英之(富山大)

- ◆ 計算機シミュレーションにおいて従来のセクタ走査法と集束ビームを用いた開口合成法により血流速度推定を行い、真値との誤差を評価する。それらの誤差を比較することにより、心臓超音波検査における開口合成法の有用性を検討する。
- ◆ 開口合成法とセクタ走査法を用いてビーム送信回数、フレーム数、角度分割数、アポダイゼーション(Tukey 窓)の係数をそれぞれ変えてシミュレーションを行い、速度推定値を算出した。推定した速度について、速度真値との誤差評価を二乗平均誤差と偏り誤差をもとに行った。
- ◆ Fig. 1は開口合成法で誤差が最も小さくなった2つの条件における結果と通常のセクタ走査法の結果を比較したものである。セクタ走査法と比べ開口合成法の誤差が小さいことが読み取れる。

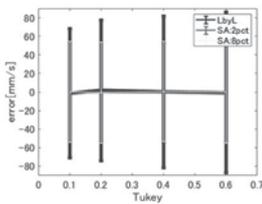


Fig. 1: Bias errors and root mean square errors of different methods. (blue) 61 angles (line by line). (red) 31 angles (SA: 2 Txs/packet). (yellow) 21 angles (SA: 8 Txs/packet).

### 2-7P-15

#### 2-7P-15 超音波による人工乳首動態の計測に関する基礎的検討

Basic study on measurement of artificial nipple dynamics by ultrasound

☆上田悠介, 茂澄倫也, 大村真朗, 長岡 亮(富山大・工).

△斎藤 哲(ビジョン株), △藤原久美子(富山大・医), 長谷川英之(富山大・工)

哺乳瓶の人工乳首の最適な形態や弾性特性を検討するためには、その動態を解析する必要がある。その解析方法として、安全性の高い超音波による計測方法を検討している。本研究では、人工乳首の動態を計測するための超音波振動子を製作し、受信信号から自己相関法を用いて変位・速度を推定する手法を検討した。

その結果、製作した超音波振動子の計測の妥当性を確認し、人工乳首を手指で押して動かした時の動態を解析することができた (Fig. 1)。今後、搾乳機使用時や実際の授乳時の動態について解析を行う。

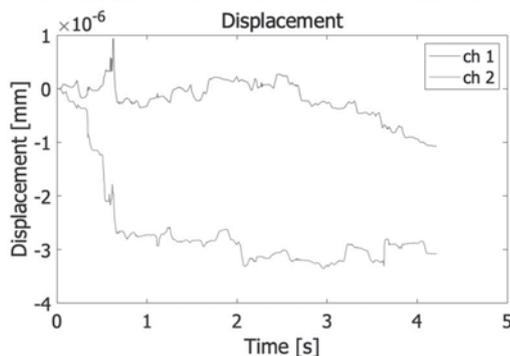


Fig. 1: Estimated displacements obtained by transducer elements, channels 1 and 2.

### 2-7P-14

#### 2-7P-14 深層学習を用いた超音波ビームフォーマの検討

Investigation of ultrasonic beamformer based-on deep learning

☆比企隆一(富山大・工), 茂澄倫也, 大村真朗, 長岡 亮, 長谷川英之(富大院・理工)

- ◆ 医用超音波画像を得るには一般的にビームフォーミング処理が必要であり、本報告では深層学習を用いたビームフォーミング処理の検討を試みた。
- ◆ 超音波イメージングファントムを使用し、エコー信号から MV 法による重みを算出した。ファントムのエコー信号を入力とし、出力がビームフォーミングに必要な重みが得られるモデルを、MV 法の重みを学習データとして畳み込みニューラルネットワーク(CNN)を学習させた。
- ◆ DAS 法 (Fig. 1)や MV 法 (Fig. 2)と比較したところ、深層学習による予測 (Fig. 3)は DAS 法と同等の画像品質であった。

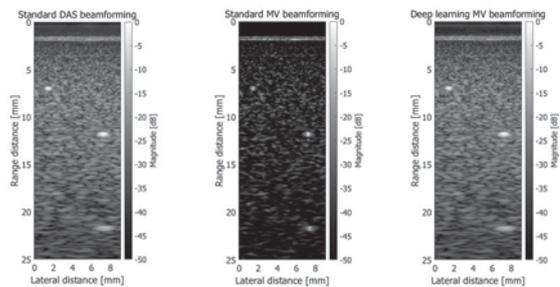


Fig.1: B-mode image of phantom obtained by DAS. Fig.2: B-mode image of phantom obtained by MV. Fig.3: B-mode image of phantom obtained by deep learning.

### 2-7P-16

#### 2-7P-16 俯角制御可能なファンビームによる港湾内の海中不審物の検出

Detection of unidentified objects in the harbor by controlling the depression angle of fan beam

☆田隈 康裕, △西野 友希, △小杉 広己, 太田 和彦(金工大・工)

港湾内外の重要施設に対して脅威を与えるような海中不審物をソーナーで探知するため (Fig.1)、指向幅の狭いサイドスキャンソーナー (Fig.2) のビーム指向範囲をサーボモーターで制御すると共に、細線化処理、Hough 変換 (Fig.3) 等の画像処理を活用し、目標の自動検出・追尾を可能とさせる機構を構築し、その有効性を海上試験 (Fig.1 参照) により確認した。

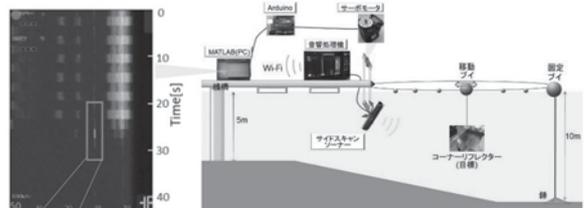


Fig.2 Output image of sidescan sonar

Fig.1 Configuration of experimental at sea

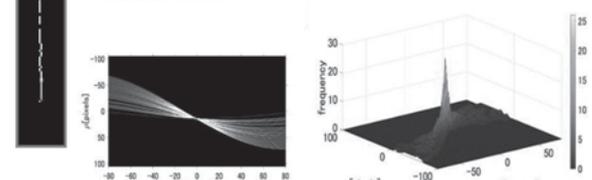


Fig.2 Target line (left) identified by Hough transform (middle, right)