

### 1-1-1

#### 1-1-1 多変量複素 Student's $t$ 分布に基づく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 based on multivariate complex Student's  $t$  distribution 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)

- ◆本稿では、独立半正定値テンソル分析 (IPSDTA) において音源生成モデルを多変量複素 Student's  $t$  分布に拡張した  $t$ -IPSDTA を提案する。
- ◆本拡張にあたり、補助関数法とベクトルワイズ座標降下法を用いて、コスト関数の単調非増加性が保証された更新式を導出する。さらに、その更新値と分布の自由度パラメータ  $\nu$  との定性的な関係を明らかにする。
- ◆実験により、 $t$ -IPSDTA の音源分離性能は ILRMA よりも高く、特に  $\nu$  が大きいとき ( $\nu \geq 10^5$ ) に高い性能を示すことが確認された。

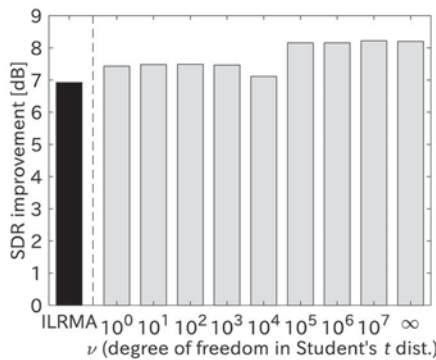


Fig.1: SDR improvement of ILRMA and  $t$ -IPSDTA.

### 1-1-3

#### 1-1-3 ランク制約付き空間共分散モデル推定を用いた多チャンネル補聴器システムの評価

Evaluation of multichannel hearing aid system using rank-constrained spatial covariance matrix estimation

☆宇根昌和(筑波大), 久保優騎(東大), 高宗典玄(東大), 北村大地(香川高専), 猿渡洋(東大), 牧野昭二(筑波大)

- ◆本研究ではスマートフォンのマイクロホンも含めた、分散マイクロホンアレー両耳多チャンネル補聴器システムを提案する。
- ◆スマートフォンを持つ人間を模したダミーヘッドを用いて、実環境下でのインパルス応答と拡散性雑音を収録した。独立低ランク行列分析 (independent low-rank matrix analysis: ILRMA) とランク制約付き空間共分散モデル推定法 (rank-constrained spatial covariance matrix (SCM) estimation) を収録したデータに適用し、分離性能の点で評価した。
- ◆スマートフォンを含めた両耳補聴器体系においても、ランク制約付き空間共分散モデル推定法は比較的高い分離性能を達成しており、特に入力 SNR が低い時に有効である。更に、目的信号の分散にスパース制約を入れた場合に、良好な分離性能を達成することを確認した。

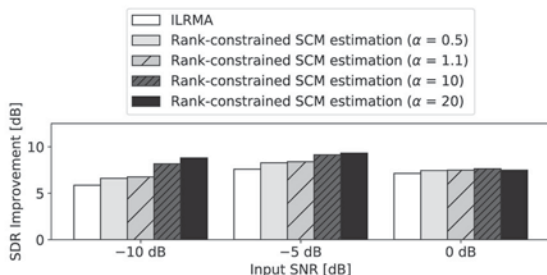


Fig.1: Separation performance for each input SNR when speaker is located at 0°.

### 1-1-2

#### 1-1-2 独立深層学習行列分析におけるマイクロホン毎の座標降下法に基づく分離行列更新

Update of demixing filter based on microphone-wise coordinate descent algorithm for independent deeply learned matrix analysis

◎牧島直輝, 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)

- ◆独立深層学習行列分析 (IDLMA) はブラインドな空間モデルの推定と DNN による音源モデルの推論を組み合わせた音源分離手法である。
- ◆IDLMA では空間モデルの推定に反復射影法 (IP) が用いられていた。IP は分離行列を音源毎 (行ベクトル毎) に推定する手法である。
- ◆本稿では、IP よりも数値的に安定な空間モデル推定法として、**分離行列をマイクロホン毎 (列ベクトル毎) に推定する手法**を提案する。
- ◆DNN を利用することで、IP 及び提案法のうち、より分離性能が向上すると期待できる方を自動選択することが可能である。
- ◆音楽信号を用いた実験により、自動選択では提案法がより多く選ばれており、従来の IDLMA から分離性能が向上することを示す。

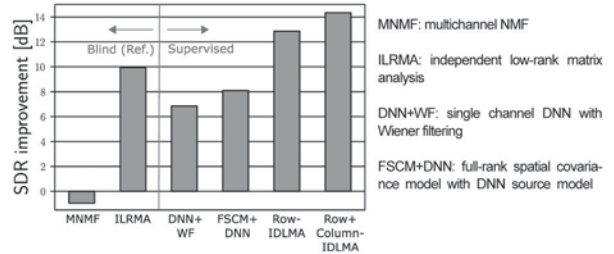


Fig.1: Average SDR improvement of 25 bass/vocal songs. Proposed "Row+Column-IDLMA" outperforms conventional "Row-IDLMA".

### 1-1-4

#### 1-1-4 楽曲音源分離のための個別音源マスク推定ネットワークの統合

Integration of Mask Estimation Networks for Individual Music Sources in Music Source Separation

☆大竹徹郎, 関翔悟, 戸田智基 (名古屋大)

- ◆楽曲信号のスペクトログラムに対して、目的音源を抽出するための時間周波数マスクを推定する手法がある。U-Net に基づく歌声分離では、大量の楽曲サンプルを用いて、歌声信号を抽出するための時間周波数マスクを推定するニューラルネットワークを学習する。歌声に限らず、個別の楽曲音源に対する U-Net も併用することで、楽曲音源分離が可能となる。一方で、このとき推定される時間周波数マスクは各音源ごとに独立に推定されるため、分離性能が制限される可能性がある。
- ◆本稿では、分離音を再度重畳したら元の混合音に戻るという再構成制約に着目し、U-Net によって個別に推定された時間周波数マスクの統合法を提案する。収録された楽曲信号を用いた実験の評価により、提案法の有効性を確認した。

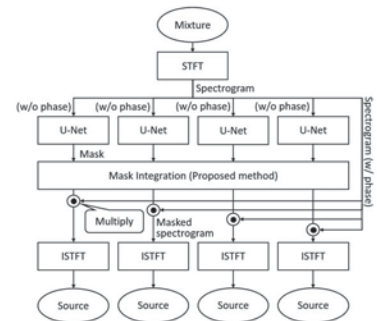


Fig.1: Overview of the proposed method

### 1-1-5

#### 1-1-5 オンライン音源分離・音源数推定・ダイア リゼーションの同時最適実現のための 深層学習モデル

All-neural online source separation, counting and diarization

○木下慶介<sup>1</sup>, ティロ フォンノイマン<sup>1,2</sup>, マーク デルクロア<sup>1</sup>, 荒木章子<sup>1</sup>,  
中谷智広<sup>1</sup>, ラインホルド ヘブウンパツハ<sup>2</sup>

<sup>1</sup>NTT, Japan <sup>2</sup>Paderborn University, Germany

- ◆従来は、別々のタスクとして扱われてきた、音源分離、話者数推定、ダイアリゼーションを、同時に最適に且つオンライン処理で実現するためのニューラルネットワークの枠組みを提案します。
- ◆複数の話者が断続的にしゃべるような、自然な会話を模したデータで、提案法の有効性を示します。

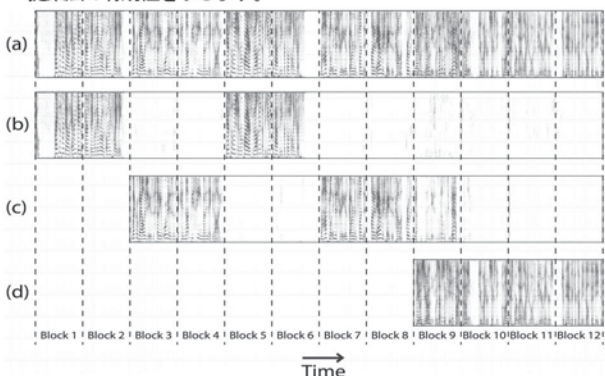


Fig. 1: Spectrograms of observed signal containing 3 speakers (a), that of extracted speech signals (from (b) to (d)). The dashed lines correspond to processing boundaries. Each color in (a) corresponds to each speaker in the observation.

### 1-1-7

#### 1-1-7 空気／体内伝導音の対応関係を活用した 自己発声音強調／抑圧法

Self-produced Speech Enhancement and Suppression Using Correspondence between Air- and Body-conducted Signals

☆高田萌絵, 関翔悟, Patrick Lumban Tobing, 戸田智基(名大)

- ◆本稿では、空気／体内伝導音の対応関係を活用した半教師あり自己発声音強調／抑圧法を提案する。従来法では、自己発声音を支配的に含む体内伝導音を利用することで強調／抑圧性能を改善した一方、分離音に体内伝導音の特性が混入する問題があった。この問題に対処するために、提案法では、空気伝導音と体内伝導音の非線形な関係性を捉えるための新たな音源モデルを導入する。提案法の有効性を調査するために、実験的評価を行い、従来法と比較して自己発声音強調および抑圧精度が改善できることを示した。

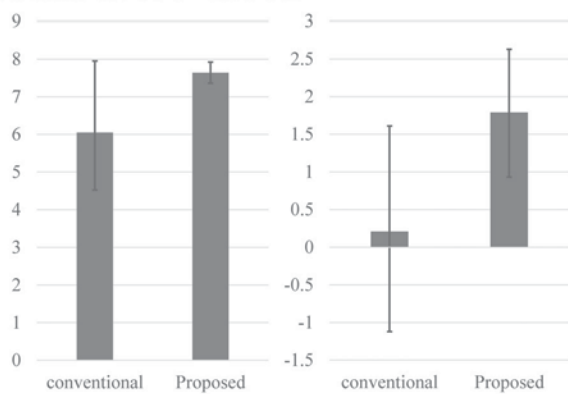


Fig. 1: SDR improvement [dB]

### 1-1-6

#### 1-1-6 Overdetermined Independent Vector Analysis: Using More Microphones at a Reasonable Computational Cost

©Robin Scheibler (TMU/JSPS), Nobutaka Ono (TMU)

- ◆We propose OverIVA, an extension of blind source separation based on independent vector analysis (IVA) to the overdetermined case with more microphones than there are sources.
- ◆We introduce a parametrization of the demixing matrix that allows to substantially reduce the number of parameters to estimate.
- ◆We combine the efficient updates of AuxIVA with orthogonal constraints between the signal and noise subspaces.
- ◆The algorithm bridges the gap between single source extraction and determined blind source separation.
- ◆Numerical experiments reveal the performance to be similar to determined separation at a fraction of the computational cost.

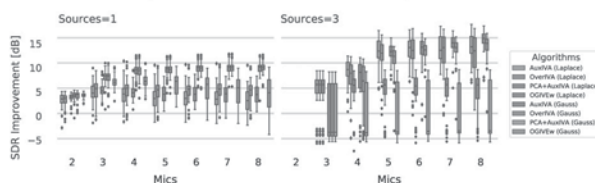


Fig. 1: Separation performance for 1 and 3 sources with up to 8 microphones in terms of signal-to-distortion ratio (SDR). We compare full AuxIVA, OverIVA, PCA followed by AuxIVA, and OGIVEw (single source). Both Gauss and Laplace IVA are evaluated.

### 1-1-8

#### 1-1-8 短時間フーリエ変換の冗長度が DNN音源強調に与える影響

Effect of the redundancy of short-time Fourier transform for DNN-based sound source enhancement

◎竹内大起, 矢田部浩平(早大理工), 小泉悠馬, 原田登(NTT), 及川靖広(早大理工)

- ◆深層ニューラルネットワーク (DNN) を用いた時間周波数マスキング  
⇒時間周波数変換として短時間フーリエ変換 (STFT) を利用
- ◆STFTには冗長度というパラメータが存在  
⇒スペクトログラムの解像度に対応するパラメータ  
⇒DNN音源強調において冗長度の設定は重要であると考えられる
- ◆本研究では、STFTの冗長度の設定がDNN音源強調の処理結果に与える影響について報告する  
⇒特定の冗長度で性能が向上することを確認

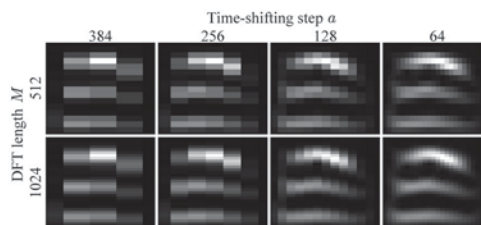


Fig. 1 様々な冗長度のスペクトログラム

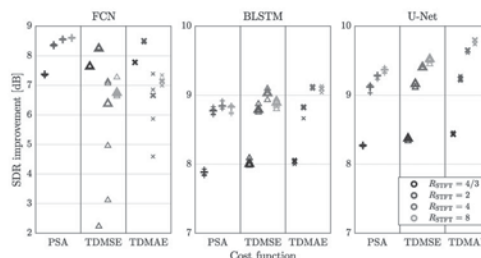


Fig. 2 冗長度とDNN音源強調によるSDR改善量の関係

### 1-1-9

#### 1-1-9 複素ニューラルネットワークによるマスク推定を用いた特定方向音声認識技術の基礎検討

Mask-based beamforming using complex-valued neural network for recognition of spatial target speech

◎早川大智, 籠嶋岳彦, 藤村浩司

- ◆複素ニューラルネットワーク (以下 CVNN) をマスク推定に適用し、特定方向から到来した音声のみを強調し認識する手法を提案する。
- ◆CVNN は、あらかじめ設定したターゲット方向から到来した音声を目的音として扱い、それ以外の方向から到来した音声および環境雑音を非目的音として扱うように学習される。
- ◆CVNN によって推定された目的音、非目的音マスクを用いて、GEV ビームフォーマで目的音を強調する。
- ◆提案手法では、ターゲット方向から到来した音声に対する認識性能が、従来のニューラルネットワークを用いた手法と同等以上の性能であること、およびそれ以外の音声、雑音は抑圧できていることを示す。
- ◆提案手法では、マイクロホンの個数未滿の方向しか強調できない従来手法と異なり、2ch のマイクロホンで2方向のターゲット方向を強調できる。

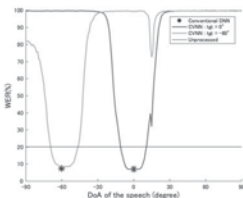


Fig. 1: WERs for each DoA of the speech (1 tgt. direction)

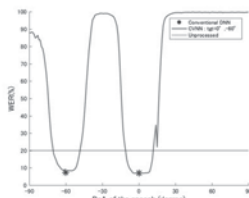


Fig. 2: WERs for each DoA of the speech (2 tgt. directions)

### 1-1-11

#### 1-1-11 音場のカーネル補間に基づくフィードバック型三次元空間能動騒音制御

Feedback spatial active noise control in 3D based on kernel-induced sound field interpolation

☆伊東勇登(東大), 小山翔一(東大, JST さきがけ), 植野夏樹, 猿渡洋(東大)

- ◆分散配置したマイクロホンとスピーカを用いて、三次元空間における領域的な騒音の低減を実現する、フィードバック型能動騒音制御(ANC)の手法を提案
- ◆従来の多点制御に基づく方法では、制御点のみの音圧パワーを最小化するため、領域的な騒音抑圧は困難
- ◆提案法では、音場のカーネル補間に基づき、領域内の連続的な音圧分布を二乗積分したパワー値をコスト関数とする
- ◆三次元音場を対象としたフィードバック型 ANC のための重み付き NLMS アルゴリズムとして定式化し、従来の多点制御法と同等の計算量で空間 ANC を実現
- ◆数値シミュレーションにより、提案法が従来法よりも高い騒音抑圧性能を示すことを確認

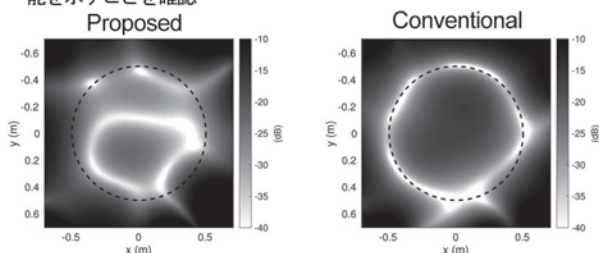


Fig. 1: Power distribution of sound field in dB after 3000 iterations when the target region was set to be circular disk.

### 1-1-10

#### 1-1-10 マイクロホンアレイとプリンキーを用いたマルチモーダル複数音声強調

Using microphone arrays and Blinkies for multi-target beamforming in reverberant environment

☆堀池大樹(首都大), シャイブラー ロビン(首都大/JSPS), 若林佑幸, 小野順貴(首都大)

- ◆マイクロホンとLEDにより音光変換を行うデバイス プリンキー(Fig. 1A)を分散配置し、ビデオカメラで撮影することによって(Fig. 1B)、広範囲の音強度情報を簡単に取得できる新しい音響信号処理の枠組みを研究している。
- ◆これまでに提案したプリンキーによる音強度分離を実際にビームフォーマの設計に応用し、マイクロホンアレイとプリンキーを用いて複数音源に対するビームフォーマを設計する枠組み(Fig. 2)を提案する。
- ◆拡散雑音下における性能評価実験を行った。実験結果(Fig. 3)からわかるように、既存のブラインド音源分離手法と比較して、同等もしくはそれ以上の性能を示した。

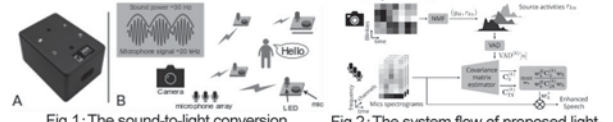


Fig. 1: The sound-to-light conversion architecture

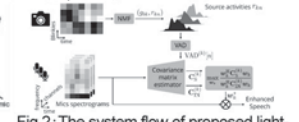


Fig. 2: The system flow of proposed light-aided Maximum SINR beamforming

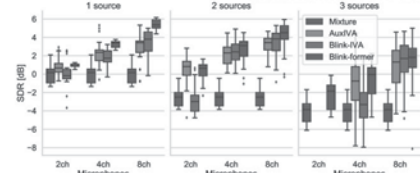


Fig. 3: The Result of the separated signals in the numerical experiment

### 1-1-12

#### 1-1-12 ゲインスケジューリング制御に基づくアクティブ音響エネルギー制御の任意境界条件音場への適用

Active control of acoustic energy in sound field with arbitrary boundary condition using gain scheduling control

☆東川晶, 鮫島俊哉(九大)

音場のアクティブ制御のロバスト性を向上させる制御手法として、ゲインスケジューリング (Gain Scheduling, 以降 GS) 制御理論を適用したアクティブ音響エネルギー制御を提案する。GS 制御は、制御系が依存する動特性が変動しているとき、その変動に合わせて制御器の特性を変動させることでロバスト性を向上させる制御手法である。閉空間音場の固有周波数は温度に依存しているため、変動するパラメータとして温度を選択することで、温度変動する音場において提案手法がロバストに音響エネルギーを共鳴周波数付近で抑圧することを示す。

Arbitrary shaped sound field modeled by FEM

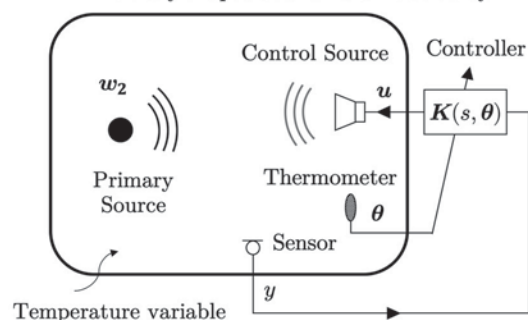


Fig.1 Configuration of GS control system.

### 1-1-13

#### 1-1-13 フィルム状圧電スピーカを用いた能動騒音制御による左右端回折音低減検討

Active noise control system for diffracting sounds around a wall using piezoelectric film speaker

○大戸康平, △河本裕介(日東電工), △牧山雄祐, 梶川嘉延(関西大)

オフィスなどの室内にて使用するパーティションを想定し、左右端回折音を低減する能動騒音制御システムについて検討した。本システムはパーティションと二次音源が一体化しており、二次音源にはフィルム状圧電素子を用いたスピーカを使用した。本システムの消音効果および消音領域の広さについて、500Hz, 800Hzの周波数で従来の平面型スピーカと比較検討した。消音領域は、制御領域内の全測定点のうち6dB低減している点の割合をRA, 0dB以上増幅している点の割合をAAとして評価した。その結果、Fig.1, Fig2に示すようにフィルム状圧電素子を用いたスピーカを二次音源として使用した場合の方がいずれの周波数においてもRAが大きかつAAが小さい結果となり、広い消音領域を得られることがわかった。これは、フィルム状圧電素子を用いたスピーカがその構造上、放射面で自由端振動モードを有しているため、回折音場とフィルムスピーカの放射音場に類似性があることが理由と考えられる。

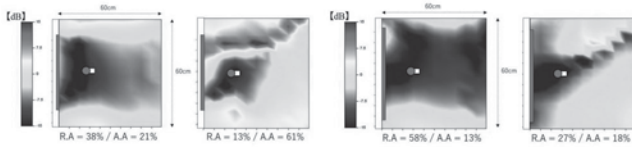


Fig.1 平面型スピーカの消音領域 左: 500Hz 右: 800Hz Fig.2 フィルム状圧電素子を用いたスピーカの消音領域 左: 500Hz 右: 800Hz

### 1-1-15

#### 1-1-15 平面複数円形アレイを用いた3次元近傍エリア再生

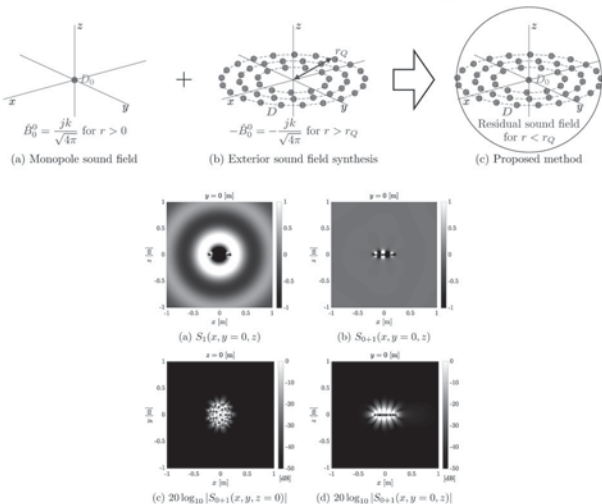
Generating 3D localized sound zone using multiple co-centered circular loudspeaker arrays

○岡本拓磨 (NICT)

Q: 「スピーカアレイの近くでのみ音が聞こえる近傍エリア再生ってできるんすか??」

オカモト: 「それならこれで!!」

- (a): 原点に配置した点音源の放射音場を
- (b): 複数の円形アレイで相殺
- (c): 差分音場による3次元近傍エリア再生を実現



### 1-1-14

#### 1-1-14 畳み込み型残留エコーパワ推定の精度向上に関する検討

Study on accuracy improvement of convolutive residual echo power estimation

○福井勝宏(NTT), 島内末廣(金沢工大), 小林和則, 原田登(NTT)

- ◆短時間スペクトル振幅 (STSA) 推定に基づくエコー抑圧処理は、音響エコーキャンセラにおける、適応フィルタ後の残留エコーを抑圧するポストフィルタとして有用である。
- ◆本報告では、エコー経路パワスペクトル推定式がステップサイズ更新式に置き換えられることに着目し、遠端側と近端側の同時通話(ダブルトーク)時に近端話者成分の影響を軽減するための、ステップサイズ制御を適用した新しいエコー経路パワスペクトル推定手法について検討した。
- ◆シミュレーション実験の結果、提案手法は、遠端側のみの通話(受話シングルトーク)時のエコー抑圧量を劣化させることなく、ダブルトーク時の音声歪み量を低減できることがわかった (Fig. 1)。

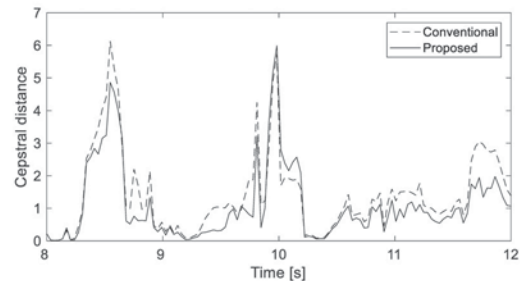


Fig.1: Comparison of LPC-cepstral distances during double-talk situation

### 1-1-16

#### 1-1-16 波数領域における時間変化に着目した二次元音場解析

Two-dimensional sound field analysis focusing on time variation in wavenumber domain

☆服部若菜, 谷川理佐子, 及川靖広 (早大理工)

- ◆光学的音響計測
  - 音の空間的な伝搬を観測可能
  - 計測結果の解析方法については十分に検討されていない
- ◆波数スペクトル解析の提案
  - 音波の振幅・周波数・進行方向を同時に可視化可能
  - 時間変化に着目 ➡ 過渡的な状態の観測が可能
- ◆正弦波・チャープ信号 (Fig. 1)・超音波の音場解析
  - 波数スペクトルの時間変化の観測を確認

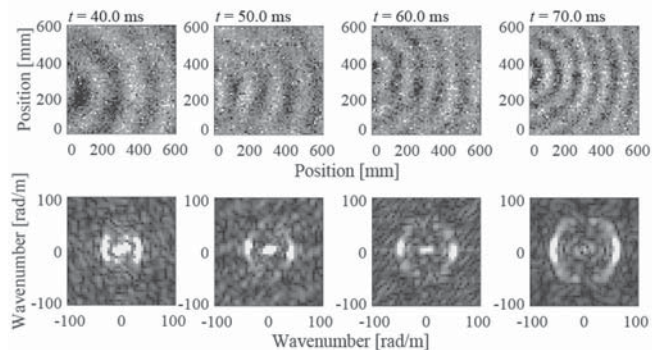


Fig.1: Wavefronts and wavenumber spectra of chirp signal

## 1-1-17

### 1-1-17 半球状超音波アレイを用いた空中音響ピンセットの大型化と保持力向上

Improvement by up-sizing mid-air acoustic tweezers using hemispherical ultrasonic transducers array

☆山本 豊(首都大院・シス), 大久保 寛(首都大院・シス)

**概要** 本研究では、我々がこれまで開発を行ってきた、空中音響ピンセットの物体保持力向上のために、半球状超音波アレイの大型化を検討し、作製したアレイの音場の計測と空中で物体を補正する実験を行い、保持力の向上を確認した。

**大型化アレイ** 大型化したアレイ(振動子数438個、径20cm)と過去の研究で作製したアレイ(振動子数80個、径10cm)のモデル図をFig.1に示す。大型化したアレイの中心では、設計通り音圧のピークが形成されていることが確認できた。計測した音場をFig.2に示す。

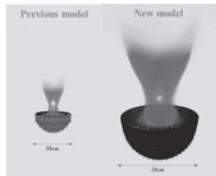


Fig.1 製作した半球状超音波振動子アレイのモデル

**物体の浮揚** 発泡ポリスチレン、木片、水滴というそれぞれ密度の異なる物体を小型超音波アレイと大型化超音波アレイを用いて浮遊を試み、比較を行った。

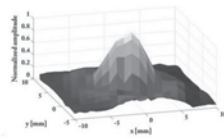


Fig.2 焦点近傍の音場

結果として、大型化アレイでは小型アレイと比較し、小さい入力電圧で物体を浮遊させることが出来、保持力の向上が確認できた。また、小型アレイでは浮遊不可能だった木片と水滴を浮遊させることが可能であった。

## 1-2-2

### 1-2-2 聴取状態の違いを考慮した車両接近報知音の開発

Development of acoustic vehicle alerting system sound considering difference of listening situation

○松田礼(日大・理工), △鈴木真里(NEC), 町田信夫(日大・理工)

- ◆車両接近報知音とはハイブリット車や電気自動車等が低速度走行時に車両前方から発音し、歩行者に車両の接近を知らせる音である。
- ◆本研究は、歩行者がはっきり聞こえる明瞭性と心地良いと感じる快適性を備え、かつ歩行者の聴取状態が異なる場合でも車両接近に気づき、安全に回避できる認知性をもつ報知音の開発を目的としている。
- ◆明瞭性と快適性を備え、歩行者が車両接近に気付いて安全に回避できる認知性と、異なる聴取状態でも同じ音条件で使用できる汎用性の高い報知音を実験的に検討した結果、条件を満たす4つの音条件を得た。

【得られた音条件】①卓越周波数I 630 Hz・卓越周波数II 3150 Hz・音圧レベル差0 dB・定常音, ②卓越周波数I 630 Hz・卓越周波数II 3150 Hz・音圧レベル差0 dB・変動周期0.5s・変調度0.45, 他2条件

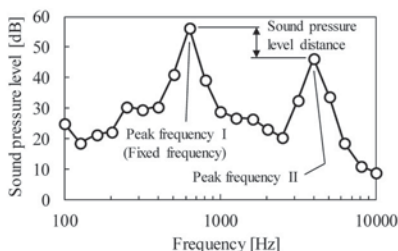


Fig.1: Example frequency spectrum of audible vehicle alerting system (AVAS) sound model

## 1-2-1

### 1-2-1 電動化・自動化時代の自動車の音デザインに向けて

Toward the sound design for electric and autonomous vehicles

○山内 勝也(九州大芸工)

- ◆ **車は変わる！音デザインも変わる！**  
自動車技術の「100年に一度の大変革期」における音デザインの課題について現状や課題を共有し、今後の展望を議論する。
- ◆ **自動車外の音**
  - 接近通報音のデザイン：国連協定規則 R138 は定められたが、依然として研究の必要性は大きい。振幅変動や聴取者の状態が検知性に与える影響は重要な課題。
  - 電動化で「どの程度」静かになるのか？：電動走行音の音源特性、特に低速度域の理解や接近通報音の影響検討が重要。ただし、道路騒音低減には駆動系騒音対策だけでは不十分。
- ◆ **自動車内の音**
  - 自動車内サイン音のデザイン：先進機能を直感的に理解できるような情報提示デザインや、自動運転の権限移譲時のサイン音デザインなど新しい課題が出てきている。自動車内でのサイン音の提示方法や音声コミュニケーションの評価も重要に。
  - 電動化による車内音環境の変容：安全性などの他性能とのバランスを俯瞰統合的にデザインするアプローチが必要。
- ◆ **渾沌の先にある音デザイン**  
機能性を満たした上で美的感性に訴える音、現存しない新たな音を創造するためには、従来にないアプローチが必要。

## 1-2-3

### 1-2-3 振幅変動に着目した気づきやすい車両接近通報音のデザインに関する検討

Design of alert sound for hybrid and electric vehicle based on amplitude fluctuation

○安井 希子(埼玉大院・理工学研), 三浦 雅展(国立音大)

- ◆国連の協定規則 R138 に定められていない振幅変動に着目し、気づきやすい接近通報音のデザインについて検討している。
- ◆用いた振幅変動の特徴下記の通りであり、Fig.1にそれらを示す。
  - 変動の速さに相当する、変動周波数
  - 不均一な変動を発生させる、時間逸脱(DT)と振幅逸脱(DA)
  - 振幅包絡
- ◆変動を付与する音はR138に基づいて合成した3通りの複合音とした。
- ◆評価実験では、録音した市街地の音を提示しながら、振幅変動を付与した刺激音を提示し、刺激音を知覚できる提示レベルを測定している。
- ◆実験結果より、
  - 振幅変動を付与することで、提示レベルが定常音よりも小さくても気づくことができる
  - ⇒振幅変動による気づきやすさ向上の効果を確認

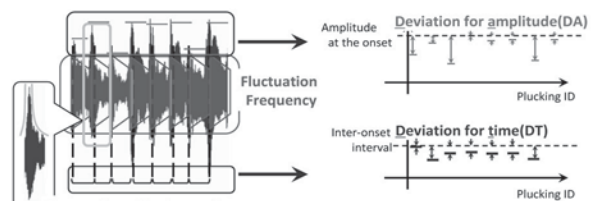


Fig.1 Outline of features for amplitude fluctuation

### 1-2-4

#### 1-2-4 自動車 HMI における情報提示音のサウンドデザインについて

Information sound design for vehicle

○鈴木広行, 平尾 章成, 中村美和(日産自動車)

- ◆高機能複雑化する自動車のHMIにおいて、ドライバーの直感的な情報の理解を助けるサウンドデザインの考え方を紹介する。
- ◆世の中の Sound HMI に対する自動車 HMI の現状を紹介。
- ◆高機能化に伴い、増え続け複雑化する機能を分かり易くするために、機能を分類化(Fig.1)。
- ◆各機能分類のイメージにマッチする。かつ、違う分類が同時に鳴った際にも、聞き分けし易くするための音色のイメージを明確化(Fig.2)。
- ◆今後期待される自動車用 Sound HMI について考察。

機能分類		代表例
警告	注意喚起	
	運転支援関連	自動ブレーキ、車線逸脱警報
	操舵支援関連	ハンドル操作警告、手放し警告
	故障/エラー関連	空気圧低下、ブレーキ異常
報知	リマインダー	シートベルト締め忘れ
	外部灯火状態	ウィンカー/ハザード
	走行状態	リバース
	操作フィードバック	エアコン操作、各機能ON/OFF操作
	システムからのメッセージ通知	ルートガイドのお知らせ
	電話着信	電話着信
	外部からの情報受信	メール受信

Fig. 1 Function classification

機能分類	音色のイメージ
運転支援関連	非常事態である事が即座に伝わる
操舵支援関連	先進性を感じさせる
故障/エラー関連	非正常感を感じる様、不協和音等によりネガティブな印象を与える。但しネガティブ過ぎて不安を抱かない様にする
走行状態	1回の走行辺りの頻度が高い為、耳あたりを良くする

Fig. 2 Image of Tone

### 1-2-5

#### 1-2-5 これからの自動車のサウンドデザインの課題

Future challenges for sound design in vehicle

○石田 康二

- ◆ Fig.1は、非自動運転車→自動運転車による人とクルマのインタラクションの変化と、それぞれが受け持つ範囲の変化を内省/機能として示す。
- ◆非自動運転時は、音は、人が操作して変化するクルマの挙動によって生じ、また他の刺激も含め、その音を聴いて操作にフィードバックするという介在の仕方をする。そのインタラクションが強いほど、音の意味的価値は高い。
- ◆自動運転時は、主に音は、様々な状態の情報を乗員に伝えるという役割が主である。音声の意味を伝える重要性は、自動運転シナリオベースの状況に依存し、人とクルマのコミュニケーションは、非自動運転時とは異なる次元の複雑さが生じる。自動運転 Level によって、自動車側から人へ伝える情報も緊急度などの状況に応じて変わるが、その時の音(ノイズ)環境下に応じた適切な聴覚情報の伝達も、サウンドデザインの出番といえる。

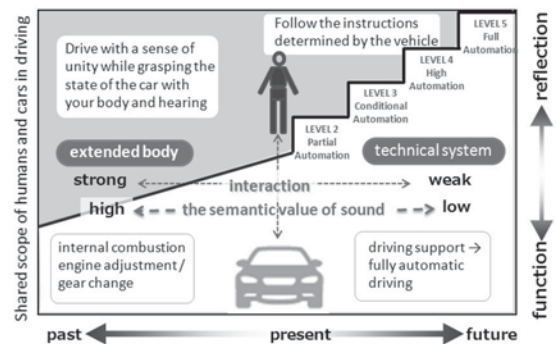


Fig.1 Changes in human-car interaction due to the evolution of vehicle

### 1-2-6

#### 1-2-6 自動運転時の権限委譲シーンを想定したサウンド UI に関する探索的検討

An Exploratory Research of Sound User Interface Design used in Take-Over-Request Situation on Autonomous Vehicles.

☆荻野朝絵, 山内勝也(九州大芸工), 浅川香, 栗野智治, 木村勝(三菱電機)

- 高度な自動運転時には、一定の条件下でシステムが自動運転を停止し、人間に車両の制御が委ねられる権限移譲 (Take-Over-Request: TOR) が行われることが想定される。適切に情報を伝えるためにサウンド UI に求められる条件を明らかにするために、本研究では音の警戒感の程度に着目して、TOR 場面での使用に適した音のデザイン手法の調査を行なった。
- まず、多種の音刺激を用意し、緊急感、及び不快感の評価実験を行い程度が広く分布するように刺激の選定を行った。選定された刺激を用いて、自動運転を意識した実験環境にて TOR 場面としての適切性の評価実験を行い、適切性と緊急感の対応やその他尺度との関連について検討を行った。
- その結果、中程度の緊急感を持つ音が TOR としての適切性の高評価につながる可能性が示唆された。また、適切性の高評価の要因として安心感、好ましさ、覚醒度などの複数の要因が影響することが示された。

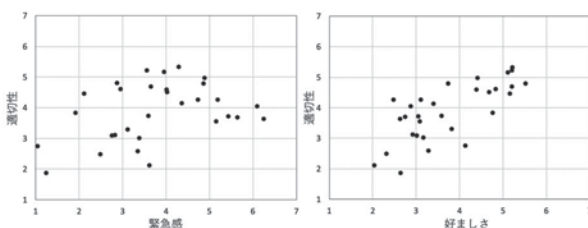


Fig. 1 適切性と緊急感の対応 (左) と適切性と好ましさの対応 (右)

### 1-2-7

#### 1-2-7 自動車室内サイン音の定位制御を用いた視覚情報確認支援における音源定位難易度の影響

Influence of difficulty of sound source localization in assistance to driver finding visual information in vehicle cockpit using localization control of sign sound.

◎中貴一, 山内勝也(九州大芸工), △田上宣昭, △川田歩(パイオニア)

- ◆自動車運転中の視覚情報確認において、サイン音の定位を制御し、視覚情報の方向を伝えることで、ドライバーは安定した運転を維持しつつ正確な視覚情報確認を行うことができる。
- ◆今回は音源定位が容易なものと困難なものの2種類を用い、視覚情報確認支援の効果調べた。
- ◆音源定位が困難な音色を用いて視覚情報の方向を伝えた場合、定位制御を用いてその方向をドライバーに伝えても一定数の見逃しが発生し、その作業負担も高いと評価された。また、視覚情報確認作業によりステアリング操作が不安定となった。
- ◆定位制御を用いた視覚情報確認作業支援を行う場合、音源定位が容易な音色を選択する必要があることを示した。



Fig.1 実験装置

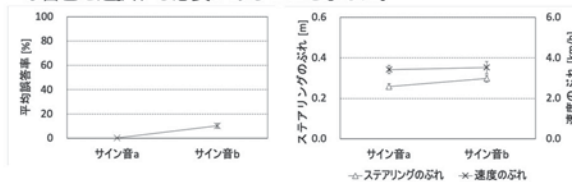


Fig. 2 音源定位が容易なサイン音 a と困難なサイン音 b における視覚情報確認できなかった割合 (左図) と運転作業のぶれ (右図)。エラーバーは標準誤差を示す。

1-2-8

1-2-8 接近通報音が電動自動車走行音の音響的特性に与える影響

Study on the effect of additional alerting sounds on pass-by noise of electric powered vehicles

☆山崎隆大, △吉野 丈, 山内勝也 (九州大芸工)

- ◆EV及びHEVを試験車とした測定データによって、接近通報音の有無による走行騒音のA特性音響パワーレベル $L_{WA}$ 、A特性相対パワースペクトル、及び周波数別の指向特性の違いを比較した。
- ◆10 km/hの場合は接近通報音により $L_{WA}$ が顕著に上昇すること、20 km/hの場合は接近通報音の影響は小さいことが確認された。 $L_{WA}$ 増加量は車種によって差があることが示唆された。
- ◆パワースペクトルの比較から、主に2 kHz周辺において接近通報音の影響が現れることがわかった。
- ◆指向特性は、車両斜め前方から側方へのレベル増加が後方と比べ大きく、スピーカ設置位置に影響を受ける可能性が示唆された。

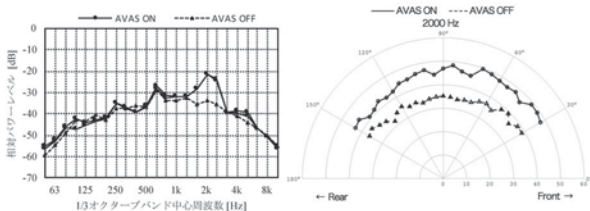


Fig. - HEV2 試験車・10 km/h 時のパワースペクトル及び指向特性

1-2-10

PESQ による ハンズフリー・インカーコミュニケーションシステムの QoE 推定

Estimating "Quality of Experience" based on PESQ for hands-free communications in car

○栗原祥子, 福井勝宏, 原田登

(日本電信電話株式会社, NTT メディアインテリジェンス研究所)

- ◆筆者らが提案する 受聴評価法は、PESQ評価と強い線形相関があり、ハンズフリー通話の QoE (ユーザ体感品質) 評価手法としての妥当性が確認されている。
- ◆ PESQによる 会話 MOS 推定の妥当性を確認するには、PESQ 評価を考慮した会話試験を実施するとともにこれまでに提案した評価法との関係性を明らかにする必要がある。
- ◆ 本稿では、会話試験と PESQ 評価との関係性、ならびに、インカーコミュニケーション評価への応用の可能性について述べる。

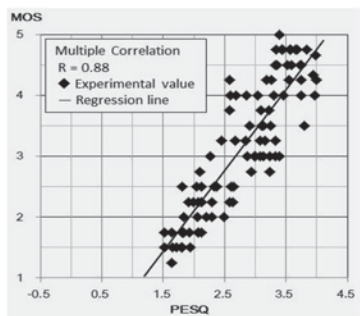


Fig.1: Relations of MOS and PESQ (◆ Experimental value, — Regression line)

1-2-9

接近通報音の検知性に二次タスクが与える影響

The Effect of Secondary Task on Detectability of Acoustic Vehicle Alerting System

☆増田真里枝, 山内勝也 (九州大芸工)

- ◆電動自動車の接近通報音の検知性は、聴取者の状態によって変わることが予想される。聴取状態の違いを二次タスクとして設定し、二次タスクの存在やそれに対する集中度が接近通報音の検知性に与える影響について検討した。
- ◆二次タスクには2種類の iPad のゲームを用意した。自動車の発進時を想定した不規則に鳴り始める音と、接近時を想定した徐々に大きくなる音を刺激として用い、背景騒音下での検知レベルを二次タスクの有無・二次タスクへの集中度で比較した。
- ◆発進時を想定した実験では、検知性に与える影響は見られなかった。接近時を想定した実験では、二次タスクによる検知レベルの上昇が確認され、二次タスクへの集中度が高いほど検知されにくくなることが示唆された。

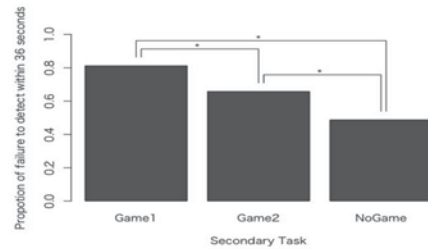


図1 音源距離が1m以下になるまでに検知できなかった割合

1-3-1

1-3-1: Using artificial speaker variability in High Variability Phonetic Training

Adriana Guevara-Rukoz<sup>1</sup>, Yutaka Yamauchi<sup>2</sup>, Alexander Martin<sup>3</sup>, Nobuaki Minematsu<sup>1</sup>,

<sup>1</sup>School of Engineering, The University of Tokyo

<sup>2</sup>School of Education, Soka University

<sup>3</sup>Centre for Language Evolution, University of Edinburgh

- Even after years of study, language learners may have difficulty perceiving L2 sounds. For instance, Japanese listeners show difficulty categorising English /r/ and /l/.
- Techniques such as High-Variability Phonetic Training (HVPT) can help improve learners' perception of this contrast.
- The methodology consists of training listeners to identify target sounds using acoustically variable stimuli, produced by a variety of speakers.
- In this work we explore the possibility of automating the creation of acoustically variable HVPT stimuli by using speech engineering tools to artificially increase speaker variability.
- To do so, we compare three types of training, differing in the number of speakers used for training: low natural variability (LNV, 1 speaker), high natural variability (HNV, 4 speakers), high artificial variability (HAV, 1 speaker + 3 artificial voices).
- We assess whether Japanese listeners trained with HNV or HAV outperform participants trained with LNV.

### 1-3-2

#### 1-3-2 自律型アンドロイド ERICA による 傾聴対話の評価

Evaluation of attentive listening dialogue with the autonomous android ERICA  
© 井上昂治, Divesh Lala, 山本賢太, 中村静, 高梨克也, 河原達也 (京大・情報学)

##### ■ 自律型アンドロイド ERICA のための傾聴対話システム

- 傾聴対話：相手の話に耳を傾けて聞く
  - 高年齢の話し相手 (傾聴ボランティア)
  - 音声対話システムで傾聴をしてくれる相手を代替できないか
- 提案システム
  - 話を理解・共感している感覚を醸し出す多様な聞き手応答
  - 相槌, 繰り返し, 掘り下げ質問, 評価応答, 語彙的応答
  - 応答例
    - 「去年の春に京都の公園に行きました」
      - 「公園ですか」 (繰り返し)
      - 「どの公園ですか」 (掘り下げ質問)
    - 「とても気持ちよく散歩ができました」
      - 「素敵ですね」 (評価応答)
      - 「そうですか」 (語彙的応答)

##### ■ 高齢者との実対話による評価

- 20名の被験者 (70代と80代で10名ずつ)
- 約5分間破綻することなく対話が継続
- 多様な聞き手応答により被験者の語りが促進された

### 1-3-4

#### 1-3-4 話速の変動を捉える特徴量に基づく留守録音声の緊急度推定

Urgency voicemail detection based on variability of speech rate  
© 神山歩相名, 安藤厚志, 増村亮, 小橋川哲, 青野裕司 (日本電信電話株式会社 NTTメディアインテリジェンス研究所)

- 話速の変動 (リズム) の変化に着目した, 留守録音声の緊急度を推定 (緊急度 高, 低の2クラス識別) する手法を提案
- 提案手法は, 1秒の窓幅を持つリズムに関連する特徴量 (長時間リズム特徴) を抽出し緊急度を推定
- 通話全体の平均話速を用いる従来手法と比べ, 誤り削減率 50%達成

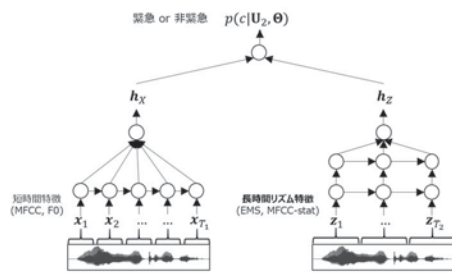


Fig. 1: Proposed neural network

	短時間特徴 (MFCC,F0)	長時間リズム特徴 (EMS,MFCC-stat)	通話全体平均特徴 (話速)	推定精度
従来手法	✓		✓	.790
提案手法	✓	✓		.895

Table 1: Accuracy of urgency estimation.

### 1-3-3

#### 1-3-3 営業電話における大規模 End-to-End 音声認識システムの活用

Telephone Sales Conversation Recognition by End-To-End Network with In-Domain Corpus Construction  
○平村健勝 (株式会社 RevComm), 篠崎隆宏 (東工大)

- 本研究では, 営業電話に着目し, 営業担当者による会話内容の振り返りや担当者間での情報共有を行うことを目的とした, End-to-End 音声認識技術の活用可能性を検証した。
- 電話営業における独特な語彙, 話し方, 音質に対応するために, PBX上で会話を録音・収集し, アノテーションを行い, 140時間を超える営業電話コーパスを作成した。
- 上記営業電話のコーパスに CSJ (日本語話し言葉コーパス)を組み合わせさせた 830 時間を超えるコーパスを用いて, WER (単語誤り率)を測定した。
- 実験には, ESPnet を用いて, Hybrid-CTC および, Transformer を用いた音声認識を行ってレシピを構築した。また, 同コーパスを用いて Kaldi (nnet3) を利用した DNN-HMM との比較を行った。
- 実験結果から, 従来方法(DNN-HMM)より, Hybrid-CTC で, 9.3 ポイント, Transformer で 6.3 ポイントの WER が改善した。
- さらに, 音声認識結果の可読性を高めるための句読点の挿入方法を提案した。
- 句読点を含めたアノテーション結果を 4.5 万件収集し, LSTM および Bi-LSTM によるニューラルネットワークで句読点挿入エンジンを作成しその評価を行い, 有用性を確認した。

### 1-3-5

#### 1-3-5 End-to-End アイヌ語音声認識における クロスリンガル話者拡張敵対学習

Cross-lingual and Speaker Adversarial Learning for End-to-End Ainu Speech Recognition  
☆松浦孝平, 上乃聖, 三村正人, 坂井信輔, 河原達也 (京大)

- 多くのアイヌ語音声アーカイブが今までに収集されているが, 大半がテキスト化されていないため, アイヌ語音声認識の検討を行う。
- 音声認識の性能は学習データの話者に依存するが, アイヌ語コーパスに含まれる話者数は少ないため, 話者に対し敵対的学習を行いモデルの話者依存性を軽減する。

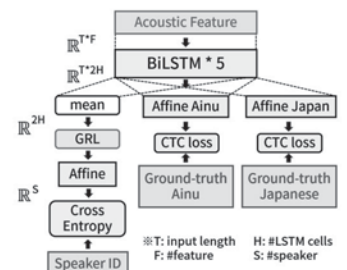


Figure 1: the proposed model

- これに加えて, 日本語コーパスを用いてクロスリンガル学習を行い, 話者数を拡張しつつ敵対的学習を行う。
- さらに, 教師無し話者適応を行う。

Table 1: ASR performance (%) for Ainu speech

	naive ASR		speaker adversarial		
	data	dev	eval	eval	
Ainu		18.4	14.9	<b>17.8</b>	<b>14.6</b>
+ Japanese		16.2	<b>13.4</b>	<b>15.4</b>	13.7
+ speaker adaptation		15.7	<b>12.9</b>	<b>14.9</b>	13.4



### 1-3-6

#### 1-3-6 Wave2Word: 音声波形を入力とする単語単位 End-to-End 音声認識

Wave2Word: end-to-end speech recognition from raw waveform to word sequence

☆上乃聖, 三村正人, 坂井信輔, 河原達也 (京大)

- ◆ End-to-End 音声認識が近年非常に高い精度を達成しているが、従来主に用いられている音響特徴量は対数メルフィルタバンクである。
- ◆ 本研究では、特徴量抽出の学習もを行い、音声波形から単語系列を直接推定する Wave2Word を提案する。
- ◆ 特徴量抽出では CNN によるフィルタを学習する。
- ◆ CSJ の評価において、従来用いられている対数メルフィルタバンクを用いた単語単位 End-to-End 音声認識モデルと比較してもほぼ同等の結果を得た。

Table 1: ASR performance (WER (%)) for CSJ-APS and CSJ-SPS.

Model	APS	SPS
40-dim lmf6 features [baseline]	12.10	9.69
Wave2Word [proposed]	11.86	9.20

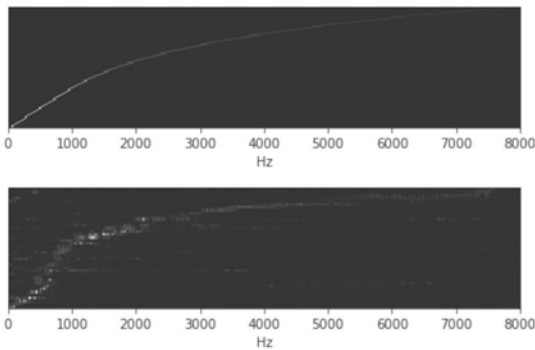


Fig.1: Learnt filter (bottom) and mel-scale filter bank (top).

### 1-3-8

#### 1-3-8 Sequence-to-Sequence model を用いた話し言葉音声認識用言語モデルのための書き言葉から話し言葉へのテキスト変換

Text Conversion of Written Language into Spoken Language for Language Model Training for Spoken Language Automatic Speech Recognition System

☆小橋優矢, 西村良太 (徳島大), 北岡教英 (豊橋技科大)

- ◆ 正確な話し言葉の音声認識には大量の話し言葉のデータが必要となる。
- ◆ 話し言葉のテキストデータを収集するのは非常に大変だが、書き言葉のテキストデータは既に多く存在するという特徴が挙げられる。
- ◆ そこで、Sequence-to-Sequence モデルを用いて書き言葉のテキストを話し言葉のテキストに変換する。
- ◆ 我々が構築したモデルではいくつかの話し言葉を非常に上手く書き言葉に変換することができた。
- ◆ また、変換前の書き言葉のデータで学習した言語モデルと変換後の話し言葉のデータで学習した言語モデルでは、後者の方が話し言葉に対してより良いパープレキシティを得ることが分かった。
- ◆ これは話し言葉音声認識にとって良い結果であると言える。

Table.1: Data example of data set using corpus

高齢者対話書き起こしコーパス		名大会話コーパス	
話し言葉 (元の文)	書き言葉 (人手で与えた文)	話し言葉 (元の文)	書き言葉 (人手で与えた文)
に、日本はきれいだけどねー	日本はきれいです	見かみはきれいな子	見かみはきれいです
けども花が好きだった	しかし、花が好きでした	ご飯が食べられるかと思って	ご飯が食べられると思います

Table.2: Perplexity of language models trained using each data

言語モデル	パープレキシティ
書き言葉 (変換前)	178.7
話し言葉 (変換後)	79.6

### 1-3-7

#### 1-3-7 共有ネットワークに基づく DNN-HMM ハイブリッド音声認識システムと End-to-End 音声認識システムの統合の検討

A study on joint modeling of DNN-HMM hybrid and End-to-End automatic speech recognition system with shared network

◎田中智大, 増村亮, 森谷崇史, 大庭隆伸, 青野裕司 (日本電信電話株式会社, NTT メディアインテリジェンス研究所)

- ◆ 共有ネットワークに基づく End-to-End 音声認識システムと DNN-HMM ハイブリッド音声認識システムの統合システムを提案
- ◆ ニューラル音響モデルの中間出力を End-to-End 音声認識システムの入力として利用することによりネットワークを共有
- ◆ DNN-HMM ハイブリッド音声認識システムの音素予測の情報を End-to-End 音声認識システムの予測に生かすことが可能
- ◆ 日本語話し言葉コーパスを用いた実験において、2つのシステムを統合することにより精度が改善することを確認
- ◆ 加えて、End-to-End 音声認識システム単体でもニューラル音響モデルの中間出力により音声認識精度が改善することを確認

Table: Character error rate (%) in different systems and input features.

System	Input features for end-to-end	%CER
(1) End-to-end	FBANK	12.95
(2) End-to-end	Internal output of AM	13.25
(3) End-to-end	FBANK + Internal output of AM	12.45
(4) DNN-HMM	-	10.85
(5) DNN-HMM + LSTM-RNNLM	-	10.17
(6) Joint (end-to-end + DNN-HMM)	FBANK	9.28
(7) Joint (end-to-end + DNN-HMM)	Internal output of AM	9.35
(8) Joint (end-to-end + DNN-HMM)	FBANK + Internal output of AM	9.18

### 1-4-1

#### 1-4-1 声帯を模した自励振動スリット下流に生成される空力音源の発生音への寄与

Contribution of aerodynamic source generated near the self-excited vibration slit to the slit sound : a vocal cord model

☆小迫誠弥, △恒吉達夫, 辻義之(名大院)

発声における発音器官である喉頭では、声帯の振動による固体伝播音や呼気流の乱れによる空力音が混在しており、具体的な音の発生要因が明らかではない。本研究では、喉頭での音発生における基礎プロセスの理解を目的として、声帯の自励振動を再現するスリットを用いて実験を行い、発生音の主要因について Lighthill-Curle の理論に基づいた考察を行った。

自励振動スリットからの発生音をマイクロホンを用いて測定した。自励振動スリットのスリット振動およびスリット下流流動場測定からマイクロホン測定点における音圧を推定し、実測音圧と比較を行った。(Fig.1 参照)。結果、スリットの振動自体に起因した音は小さく、スリット下流流動場の空力音がスリット発生音の主要因であることがわかった。また、音源の分布から空力音源はスリット下流近傍に二又状に集中して発生することがわかった。

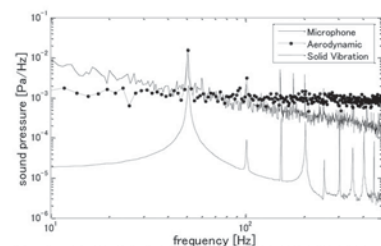


Fig.1: Measured sound pressure and estimated sound pressure

### 1-4-2

#### 1-4-2 幼児期前期におけるつくり歌の韻律と旋律

Prosody and melody of improvisational song in early infancy

○平田嘉之(名古屋市立大院・人間文化研)

Fig.1に1歳11か月女児が即興的に歌ったつくり歌の冒頭の採譜例を示す。Praatによりピッチ輪廓を可視化したところ、Fig.2に示すように、下記の4点が明らかになった。



Fig.1 Notation example of improvisational song of 1 year 11 months old girl.

(1)ピッチが安定して平坦(Flat)に伸びて音程が感じられ

るとともに、語尾のピッチが上昇(Rising)している歌の要素が多く見

られた。(2)一方で、ピッチが平坦に安定せず、語尾のピッチが下降(Falling)している言葉(語り)の要素も混在していた。(3)息継ぎ(Breath)が単語、文節、文の切れ目で行われていて、それによりフレーズを構成していた。(4)言葉の韻律と離れたメロディをつくっていた。また、このサンプルの場合、その音階的特徴により、わらべ歌的な旋律を示した。

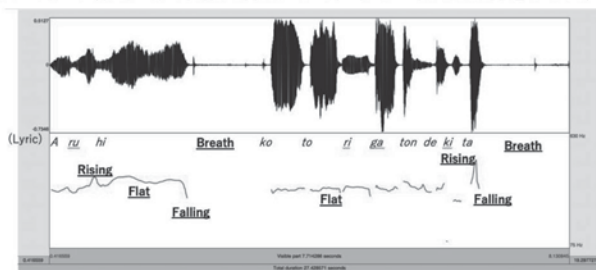


Fig.2 Pulse and pitch contours of "Aruhi kotoriga tonde kita"

### 1-4-4

#### 1-4-4 CNNに基づく歌声合成における計算量削減の検討

Computational complexity reduction method for CNN-based singing voice synthesis

◎中村和寛<sup>1</sup>, 高木信二<sup>1,2</sup>, 橋本佳<sup>1,2</sup>, 大浦圭一郎<sup>1,2</sup>, 南角吉彦<sup>2</sup>, 徳田恵一<sup>1,2</sup> (1テクノスピーチ, 2名工大)

◆目的: CNN 歌声合成における計算量の削減

- 歌声は多様な表現を含む ⇒ 時間による変動を CNN でモデル化
- モデルに再帰構造が無い, 合成時に MLPG が不要 ⇒ 高速

◆手法: モデルを異なる時間単位で駆動

- feed-forward 部分を状態単位で駆動することで計算量を削減
- 畳み込み部分をフレーム単位で駆動することで高い自然性を維持

◆結果: 少ない計算量で自然性の高い音声合成可能

- 3サイズのCNNを従来法 (FFNN+MLPG) と比較

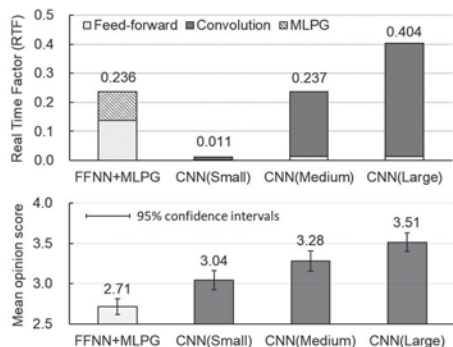


Fig1: Computational complexity and MOS result.

### 1-4-3

#### 1-4-3 ユーザ歌唱のための generative moment matching network に基づく neural double-tracking

Neural double-tracking based on generative moment matching network for users' singing

☆田丸浩気, 齋藤佑樹, 高道慎之介, 郡山知樹, 猿渡洋 (東大院・情報理工)

◆double-tracking (DT)

- 同一フレーズを多重録音し, 歌声に厚み・豊かさを付与する手法
- 2回なるべく同じように歌う必要あり → 難しく, 手間

◆artificial double-tracking (ADT): 従来法

- 原音のピッチ系列を信号処理的に変調して原音にミックス
- 発話間変動を考慮せず → 自然な多重録音感を再現できない

◆neural double-tracking (NDT): 提案法

- あたかも再度歌ったかのように歌声を変調し, 原音にミックス
- 繰り返し歌唱データを用い, generative moment matching network でピッチ系列の発話間変動をモデル化 → 自然な多重録音感!
- 合成歌声を入力とする NDT [田丸+, 音講論 (春), 2019] を自然歌声入力に拡張したもの

◆主観評価実験により, 以下を確認

- (1) NDT は ADT よりも高い多重録音感を実現
- (2) 不特定歌唱者モデルでも, 歌唱者依存モデルと同等の性能

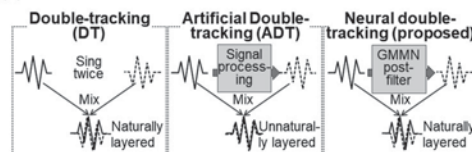


Fig.1: Double-tracking (DT), artificial double-tracking (ADT) and proposed neural double-tracking (NDT).

### 1-4-5

#### 1-4-5 日本語アナウンサによるニュース音声の文間に位置する休止

Pauses located between sentences in news speech by Japanese newscasters

◎中村 静(京大・情報学), 石井カロス寿憲(ATR・石黒研), 河原達也(京大・情報学)

◆合成音声によるアナウンサのニュース音声の自然性向上を目指して, 日本語アナウンサによるニュース音声の文間に位置する休止(文間休止)の特徴の解明とその数理的な記述を試みた。

◆文間休止持続時間の特徴の分析では, A) 直前の文の最後6モーラの平均F0, B) 直後の文に含まれるモーラの数と相関があることが明らかとなった。Aは, より低いF0が, より大きな境界, すなわち, より長い文間休止持続時間が後続することを, Bは, より長い後続文の準備のために, より長い文間休止持続時間が必要であることを示唆している。

◆これら2変数による文間休止持続時間を予測するモデルを構築したところ, 予測・実測値の相関係数は, テストデータに対して0.45であった。この結果は, 当該文間休止の前後の文や部分の特徴を反映すると考えられる多数の韻律的要素を変数とするモデルより高精度であった。文間休止持続時間との関係を考慮して, 慎重に変数を選択する必要があることを示している。

◆構築したモデルの予測精度は実用的なものではないが, 前後の発話の韻律情報から, 文間休止持続時間をある程度予測できることがわかった。今後は, 予測精度の向上のために, ニュース原稿の性質や統語的要素も考慮する。

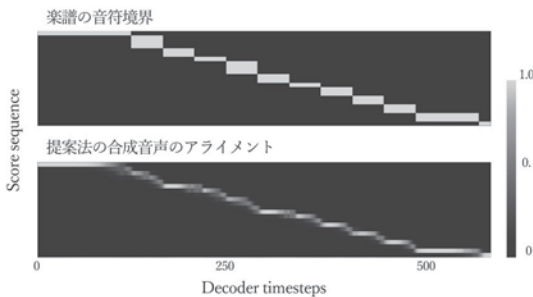
### 1-4-6

#### 1-4-6 楽譜時間情報を用いた アテンション機構に基づく歌声合成の検討

A study on singing voice synthesis with attention mechanism using musical score time information

☆村田舜馬, 藤本崇人, 法野行哉, 高木信二,  
橋本圭, 大浦圭一郎, 南角吉彦, 徳田恵一(名工大)

統計モデルに基づく歌声合成の手法の一つとして、深層構造を持つニューラルネットワーク(Deep Neural Network; DNN)に基づく歌声合成が提案されている。一方、近年では音声合成分野において、Tacotron 2などの、アテンション機構を含む sequence-to-sequence モデルにより音響特徴量と時間構造を同時にモデル化する手法が提案されている。そこで、本研究では、楽譜時間情報をアテンション機構に入力することで楽譜通りのタイミングで歌声を合成する手法を提案し、その有効性を検証する。実験結果から、提案法が楽譜通りのタイミングで発声される歌声の合成を実現した。また、約 60 分と少量な学習データにも関わらず、提案法が音響特徴量系列と楽譜情報系列の対応関係を学習可能であることを確認した。



### 1-4-8

#### 1-4-8 読み仮名と韻律記号を入力とする end-to-end 音声合成における 韻律記号の有効性

Effectiveness of prosodic symbols in Japanese end-to-end speech synthesis that inputting kana and prosodic symbols

◎栗原清, 清山信正, △熊野正(NHK 技研)

- ◆読み仮名と韻律記号を入力とする end-to-end 音声合成における韻律記号の有効性を確認するため一対比較法で自然性を評価
- ◆各韻律記号を除いた入力データでモデルを学習・合成した音声で実験
- ◆実験の結果 (Fig. 1)、本手法の入力として重要な韻律記号は「アクセント上昇記号」と「アクセント下降記号」であることがわかった
- ◆既報告で提案した韻律記号を用いる事により簡易なアノテーションで音声合成を実現できると考えられる

- ◆ 韻律記号を全て使用
- アクセント句区切り記号を除く
- × フレーズ句区切り記号を除く
- ▲ アクセント句区切り・フレーズ句区切り記号を除く
- アクセント上昇記号を除く
- ✕ アクセント下降記号を除く



Fig 1 Thurstone Case-VI尺度値 (自然性)

### 1-4-7

#### 1-4-7 End-to-end 韻律推定に向けた DNN 音響モデルに基づく subword 分割

Subword tokenization considering DNN acoustic model likelihood towards end-to-end prosody generation

☆阿曾真至, 高道慎之介, 高宗典玄, 猿渡洋(東大院・情報理工)

- ◆テキストから韻律を直接推定する end-to-end 韻律推定は、専門的な言語知識なしに韻律を生成できるため、注目されている。
- ◆これまでに、我々は end-to-end 韻律推定に向けて、subword 系列を隠れ変数、連続F<sub>0</sub>系列を観測データとした、expectation-maximization (EM) アルゴリズムによる deep neural network (DNN) 音響モデル学習を提案した [阿曾他 2019 春]。
- ◆本稿では、EM アルゴリズムによる DNN 音響モデルの学習 [阿曾他 2019 春] と、その DNN の推定精度に基づく subword の削除を反復することで、音響モデル尤度に基づく subword vocabulary を構築する。これにより音響モデル尤度に基づく subword 分割が可能になる。
- ◆実験では、(1) 提案法によるモデル尤度の改善と、(2) 提案法の分割がアクセント句を考慮している点を示した

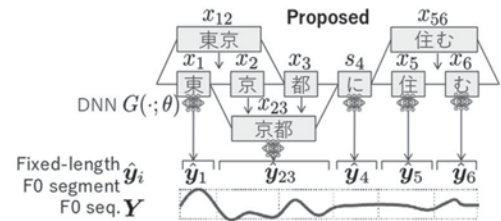


Fig.1: Conventional method and proposed method

### 1-4-9

#### 1-4-9 End-to-End 型テキスト音声合成における WaveNet ボコーダの学習に関する調査

An investigation on training of WaveNet vocoder in end-to-end text-to-speech.

☆安原和輝, 林知樹, 戸田智基 (名大)

- ◆本研究では、End-to-End 型テキスト音声合成の品質に寄与する要因を明らかにすることを目的とし、Tacotron 2における WaveNet ボコーダの学習に関して、複数の観点から調査を行う。
- ◆第一に、学習時と合成時における音響特徴量のミスマッチによる影響を緩和するため、Tacotron 2 で生成された推定音響特徴量を用いたファインチューニングを適用する。
- ◆第二に、ミスマッチに対する頑健性を高めるため、音響特徴量系列に対する時間方向の畳み込み層を適用する。これに加えて、ファインチューニングを組み合わせた方法についても検討する。
- ◆実験的評価により、以下のことが示された。

- 単なるファインチューニングの適用では自然性の改善が見られない
- 音響特徴量系列に対する時間方向の畳み込み層の適用で自然性が改善する
- ファインチューニングと畳み込み層を併用することでさらなる自然性の改善が見込める

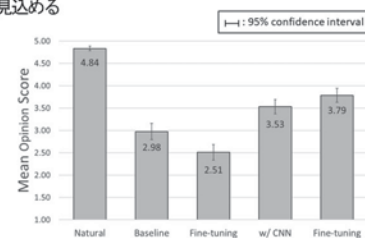


Fig.1 Result of Mean Opinion Score test

### 1-4-10

#### ソフトアテンションを使用しないエンドツーエンド音声合成の初期検討

Toward end-to-end text-to-speech synthesis without soft attention

☆安田裕介(総研大/NII), 山岸順一(総研大/NII), Xin Wang(NII)

エンドツーエンド音声合成はテキストから直接音響特徴量に変換する音声合成手法です。今まで提案されたエンドツーエンド音声合成手法はいずれもソフトアテンションを用いています。この発表ではソフトアテンションを使わない新しい手法を紹介します。

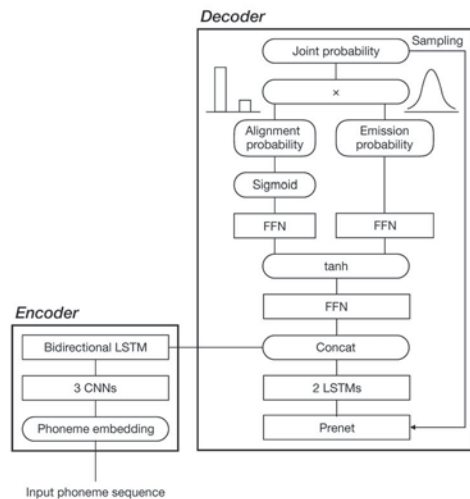


図1 提案手法のネットワーク構造。

### 1-5-1

#### 1-5-1 日本語文字列の印象に関する研究

Impressions of Japanese character strings

☆平出勇太, 山田真司(金沢工大院)

商品やキャラクターに命名する際の名前は語感が良くなるような文字列で構成されている。この語感に関して、様々な研究が行われてきたが、一字音節や、擬音語、オノマトペを対象に実験を行っていたため、普段我々が用いている文字列の語感とは違いものである。そこで本研究では、無意味文字列の印象評定実験を行い、文字列の構成に含まれる要素が、文字列の印象にどのように影響しているのかを明らかにした。

実験の結果、文字列の印象は「快さ(pleasantness)」と「シャープさ(sharpness)」の2次元の印象平面で表せることが明らかになった。さらに、文字列の印象は1音目の子音に強く影響されることが明らかになり、印象平面上に子音を布置したものをFig 1に示す。また、「快さ」は文字列内の濁音の有無から強く影響を受け、「シャープさ」は文字列の1音目の調音部位から特に大きな影響を受けることが明らかになった。



Fig. 1. Centroid of the impressions for each consonant in the first mora

### 1-4-11

#### 1-4-11 落語音声合成モデルの頑健な学習方法と発話様式の変動への対処

Robust training method for rakugo speech synthesis models and dealing with various speaking styles

☆加藤集平, 安田裕介(総研大/NII), Xin Wang,

Erica Cooper(NII), 高木信二(名工大), 山岸順一(NII)

#### 背景

- 人を楽しませる音声合成として、落語音声合成を開発している
- Tacotron でモデルの end-to-end 化を試みたが、うまく合成できなかった
- 落語は表現豊かで、発話様式が様々に変動する

#### 目的

- 落語音声合成の end-to-end モデルを頑健に学習する
- 発話様式の変動を表現できるモデルにする

#### 解決方法

End-to-end モデルの頑健な学習

- SSNT に基づく音声合成(1-4-10)を導入する
- ソフトアテンションを用いないためアラインメントを頑健に学習

発話様式の変動への対処

- Tacotron のために提案された global style tokens (GST) を、SSNT モデルに導入

#### 結果

- アラインメント誤り率が劇的に低下
- GST により役の区別・内容理解度が向上

### 1-5-2

#### 1-5-2 聴取による話者識別における言語の影響 —日本語母語話者が発話した日本語・英語の場合—

The effects of language on speaker identification

-in case of Japanese and English spoken by Japanese native speakers-

☆井上峻河(上智大), 網野加苗(科警研), 荒井隆行(上智大)

◆日本語母語話者が日本語と英語を話した場合に、他の日本語母語話者がこれらの音声聴いた際、発話に対する話者識別の精度はどのようになるのかを調査

◆話者: 非母国生(N)と母国生(R)からそれぞれ4名(計8名)

◆単語: 日本語(J)と英語(E)からそれぞれ4語ずつ(計8語)

◆音声選定: 上の話者を各2名、単語を各2語ずつ(計4名, 4語)

◆実験方法: それぞれの話者による日本語と英語の単語の組み合わせを、総当たりでAX法にてランダムに再生

◆結果: (条件ごとの正答率)「非母語である英語 > 母語である日本語」



Fig. 1: Accuracy rate of Speaker Conditions NN



Fig. 2: Accuracy rate of Speaker Conditions NR



Fig. 3: Accuracy rate of Speaker Conditions RN

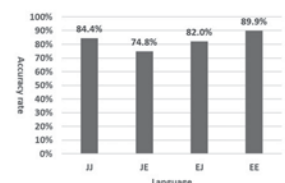


Fig. 4: Accuracy rate of Speaker Conditions RR

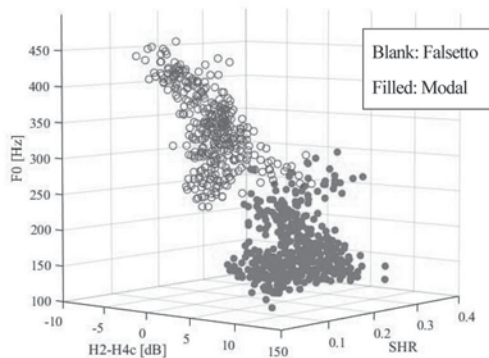
## 1-5-3

## 1-5-3 偽装音声としての裏声と通常発声の音響的判別

Acoustic discrimination between falsetto used as disguised voice and modal voice

○網野加苗(科警研), 荒井隆行, △庄野植(上智大)

- ◆法科学では裏声による音声偽装がしばしば見られる
- ◆未知話者の音声について裏声/通常発声を音響的に判別したい
- ◆資料: 成人男性 15 名による短文; 偽装を目的とした裏声と通常発声
- ◆音響分析結果: 裏声-通常発声間の差が大きかった特徴量 (下図)
  - H2-H4c 2倍音と4倍音の振幅差(補正值)
  - SHR 低調波と調波の振幅比
  - F0 基本周波数
- ◆判別実験結果: SVM, 交差検証法で99.9%の正解率 ( $n = 810$ )



## 1-5-5

## 1-5-5 小児の構音障害と支援

○平島ユイ子(福岡国際医療福祉大・言語聴覚)

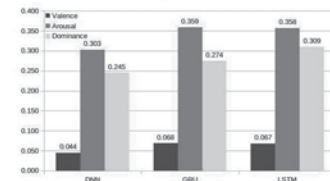
- ◆通常発達では子どもは小学校に入学する頃には誰にでも伝わる発音で話すことができるようになる。身体疾患や知的障害がないのに発音が不明瞭な場合があり、機能性構音障害と言う。語音の産生では、発話すべき語を選択した後、語を構成する音とその順序を決定し、構音器官を動かして産出に至る。従って、この過程のいずれかに問題があれば発話が不明瞭になる。発話の不明瞭さには異なる原因が考えられるため、個に応じた指導が求められる。
- ◆音声の知覚の発達は、母音は生後6ヶ月頃、子音は10ヶ月頃に知覚できるようになり、11ヶ月頃には母語の音韻体系に合致した音韻知覚を示すようになる。音韻の位置記憶や操作能力は4~7歳頃に発達する。
- ◆音韻産生の発達は、生後すぐの叫喚発声、2ヶ月頃のクーイング、6ヶ月頃の喃語の産出、10ヶ月頃のジャーゴンを経て、初語に至る。その間に母乳を飲むのに適した口腔形態から物を咀嚼し食べることができるまでに口腔機能の発達が見られる。
- ◆構音訓練として聴覚刺激法や構音位置づけ法等が用いられる。音韻聴取力を高め、構音方法を指導する。側音化や口蓋化の異常音の弁別には経験が必要である。
- ◆構音の練習は短時間であっても毎日実施する必要があるが、子どもにとっては苦手な口腔運動であり、単調な練習であるため継続しにくい。練習方法を子どもに合わせて工夫する必要がある。

## 1-5-4

## 1-5-4 RNN-based Dimensional Speech Emotion Recognition

○ Bagus Tris Atmaja, Reda Elbarougy, Masato Akagi (JAIST)

- ◆ A speech emotion recognition system based on recurrent neural networks is developed using long short-term memory networks.
- ◆ Two of acoustic feature sets are evaluated: 31 Features (3 time-domain features, 5 frequency-domain features, 13 MFCCs, 5 F0s, and 5 Harmonics) and eGeMaps feature set (23 features).
- ◆ To evaluate the performance, some metrics are used i.e. mean squared error (MSE), mean absolute percentage error (MAPE), mean absolute error (MAE) and concordance correlation coefficient (CCC). Among those metrics, CCC is main focus as it is used by other researchers.
- ◆ The developed system used multi-task learning to maximize arousal, valence, and dominance at the same time using CCC loss (1 - CCC) by the following formula.
 
$$CCCL_{tot} = \alpha CCCL_{val} + \beta CCCL_{aro} + \gamma CCCL_{dom}$$
- ◆ The result shows using LSTM networks improve the CCC score compared to baseline dense system. The best CCC score is obtained on arousal followed by dominance and valence.



## 1-5-6

## 1-5-6 言語通級指導教室の発音指導を支援する～教員・児童・保護者を支援する技術

To support pronunciation instruction in special education classes for language-disabled children

-Technology to support teachers, children and their parents

○勝瀬郁代(近畿大・産業理工)

- ◆言語通級指導教室で行われている発音指導を支援するためにこれまで開発したシステムを、背景にある教育現場の状況を交えて紹介する。
- ◆はじめは、家庭での発音の反復練習を支援するためのアプリケーションの開発から始まった。
- ◆次に、主に言語通級指導教室の教員の支援を目的として、教員と外部専門家の連携を促進するために、発音練習システムをwebアプリケーション化した。
- ◆さらに、外部専門家が、遠隔からインタラクティブに児童の構音検査を実施できる遠隔構音検査システムを開発した。
- ◆現在、発音指導に効果的な、コンピュータによる発音の段階的評価を実現すべく研究を進めている。
- ◆発音の異常は、音韻の弁別や記憶、構音運動の問題など、様々な原因によって起こっているため、障害のある子どもの発音の評価には、分野横断的な知見が必要となる。
- ◆研究分野の垣根を超えた音声コミュニケーション研究が望まれる。

## 1-5-7

### 音声の個人性に関する多角的研究 山岸順一 (国立情報学研究所)

- 音声は手軽で、自然で、直感的なモダリティの一つである。また同時に、音声の個人性は我々の生体情報やアイデンティティの一部でもあり、生体認証、音声情報処理、プライバシーなど複数の異なる分野において重要なファクターである。
- 我々は現在、日仏共同 CREST プロジェクトを通して、これらの分野にまたがる多角的研究を行っており、本講演では近年の研究成果について紹介する。
- 融合研究の基礎技術となっているニューラル波形モデルについて紹介したのち、生体認証との融合研究、プライバシー分野における融合研究について紹介する。

## 1-5-9

### 1-5-9 顔面の動きをフィードバックとして用いる 発話訓練システムの検討

Speech training system using facial movement as feedback

○北村達也(甲南大) 川村直子(姫獨大) 能田由紀子(ATR-P)  
吐師道子(県広大)

- ◆背景：顎運動制約法を用いた発話訓練では下顎以外の部分を大きく動かすこと、大きな声を出すことなどが重要である。
- ◆提案する発話訓練システム：オプティカルフローを用いて顔面の動きを計測し、リアルタイムでフィードバックする。声の大きさについてもフィードバックする。
- ◆方法：提案法により成人14名に3分間発話訓練させた。その前後の音声を収録するとともに発話がどの程度うまくできたかに関する Visual analog scale (VAS) を回答させた。
- ◆結果：
  - 訓練後に音声の振幅、基本周波数の変動幅と平均値が増加した。
  - 訓練後に VAS も上昇し、訓練に対して肯定的な意見が聞かれた。



Fig. 1: Screenshot of the proposed speech training system.

## 1-5-8

### 1-5-8 ソーシャル・イメージングによる社会的インタラクションの支援

Supporting Social Interaction by Social Imaging Technology

○鈴木健嗣(筑波大・システム情報)

- ◆ソーシャル・イメージングとは、人々の間にある社会的行動や交流状況、さらにその社会的な関係を顕在化して明示するための技術である。装着型機器や画像計測技術を組み合わせることで、人々の社会的行動を顕在化し、視聴覚を通じて本人や支援者に対し適切な先行刺激や後続刺激として明示（フィードバック）する技術といえる。
- ◆我々はこれまで、複数の基本的な行動の顕在化を実現するため、生体信号に基づく表情識別や、行動を計量する装着型機器の開発を行ってきた。また一方、環境の構造化技術として大型床面プロジェクションマッピングシステムを備えた新しい体育館における活動支援に関する研究として「ミライの体育館」という取り組みを行ってきた。
- ◆このような人の五感に関する人間情報技術は急速に発展を続けており、人の感覚系への刺激を機械系を通じた信号入力で補う情報システムが発展している。このような技術は、実世界の環境において周囲の物体や人々との積極的な相互作用を可能にしながら、人の身体を変容することで周囲環境や他者に対する自己の認識、行動、相互作用を変化させたりすることが可能になると考えている。
- ◆本講演では、これら一連の取り組みを紹介するとともに対話・会話といった言語・非言語的なインタラクションを含むソーシャル・イメージングの目指すものについて議論する。

## 1-5-10

### 1-5-10 発声学習支援のための対話的 実時間音声属性可視化ツールについて

Interactive and real-time visualization tools of voice attributes  
for learners of vocalization

○河原英紀(和歌山大), 榊原健一(北海道医療大)

羽石英里, △萩原かおり(昭和音大)

- ◆適切な発声方法を身につけることは生活の質を維持する上で重要。
- ◆具体的にイメージし難い発声の様子を実時間で可視化し対話的に操作するツールを提供することにより、発声の学習支援を狙っている。
- ◆MATLAB で実装したツール群を GitHub で公開している。

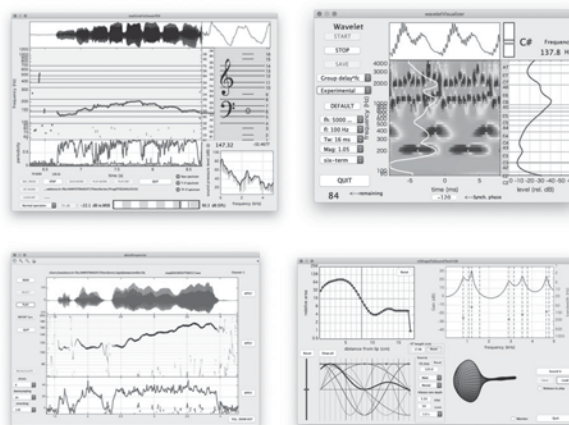


Fig. 1: Snapshots of interactive and real-time tools. (Top left) Fundamental component and spectrum, (Top right) Voice source, (Bottom left) Detailed inspector, and (Bottom right) Voice production simulator.

### 1-5-11

#### 1-5-11 音声の録音の基本手順と支援ツールについて

Protocol and support tool for recording speech signals

○榊原健一(北海道医療大), 河原英紀(和歌山大), 水町光徳(九工大)

- ◆本発表では、高品質でなくても、できる限り高い利用率を実現するために必要な、現実的な録音環境や機器に関する注意点について概説し、実装した支援ツールに関して録音の基本手順、録音環境の記録の方法を周波数領域 velvet noise (FVN)を用いた録音環境記録および音圧校正のためのツールの概要を説明する。

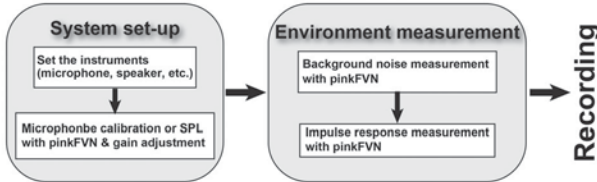


Fig.1: Schematic diagram of recording of the speech sound using our proposed tool

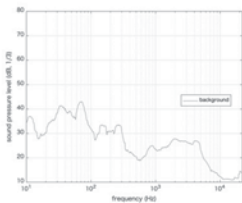


Fig.2: Example of background noise

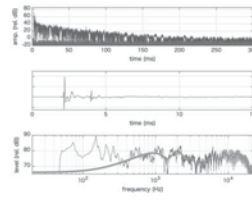


Fig.3: Example of impulse response

### 1-5-13

#### 1-5-13 無発声音声認識における話者の違いが認識に与える影響の検討

A study on the effect of the difference of speaker on silent speech recognition

◎大田健紘(日本工大・基幹工)

- ◆生体情報を用いた音声認識に取り組んでおり、生体情報として筋電図と皮膚振動を用いることの有効性を示し、下口輪筋や口角下制筋といった口を引き下げる動作に関係する筋肉が認識に有効であった。
- ◆本稿では、認識対象以外の発話者のデータも用いて音響モデルを作成した場合の認識率への影響を調査した。
- ◆本稿では、筋電図は唇の周りの筋肉である上口輪筋、下口輪筋を計測した。皮膚振動は上唇で計測した。
- ◆Figure1より、認識対象の話者のデータを全く使わずにモデルを学習した場合、認識精度は20%以下であった。モデル学習用のデータに認識対象話者のデータを40個追加した段階で70%程度の認識精度となった。
- ◆無発声音声認識においても多くの話者のデータを集めることで未知話者に対する認識精度を向上できることが示唆された。

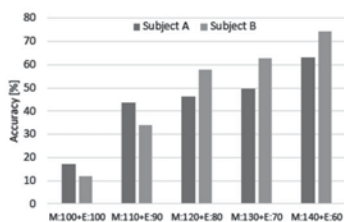


Fig.1: SER of Subject A

### 1-5-12

#### Deep Neural Network を用いた病的音声の声質評価

Evaluation of Pathological Voice Quality using Deep Neural Network

☆日高駿介, 李庸學, 若宮幸平, △中川尚志, 鍋木時彦(九州大)

- 病的音声の声質の聴覚心理的評価は音声障害の臨床において広く用いられているが、評者間変動や評者内変動が原因となり、再現性に欠ける問題がある。本研究は、この問題点を解決するため、嘔声の評価法として国内外で使用されている GRBAS 尺度の評点を出力する Deep Neural Network (DNN) の構築を試みるものである。
- GRBAS 各項目について、Bidirectional GRU と全結合層から構成される DNN を構築した。入力特徴量として、持続母音/a/の起声から終声までの対数振幅スペクトログラムと位相変化スペクトログラムを比較した。位相変化スペクトログラムは位相スペクトログラムの各周波数ビンにおけるフレーム間変化に着目した特徴量である。
- 1名の耳鼻咽喉科医により、正解ラベルを付与しデータセットを作成した。3118 サンプルを含むデータセットに対して5分割交差検証を行なった結果、信頼性を示す指標である Cohen の線形重み付けカッパ係数は、Landis と Koch の目安において、項目 B で Substantial, G で Moderate, RAS で Fair な性能を示した。項目 GRBS については、位相変化スペクトログラムを用いたときに、より高い推定精度が得られた。この結果は、複素スペクトログラムの位相成分に嘔声の声質に関連する情報が含まれていることを示唆している。位相変化スペクトログラムが持つ意味の検証や推定精度のさらなる向上、および臨床で医療従事者が利用するためのアプリケーション開発が今後の課題である。

### 1-5-14

#### 1-5-14 構音障害を支援するウェアラブルな音声補完システムの簡易声質維持機能の試作

A prototype of a simple voice quality reproduction function on a wearable voice correction system that supports articulation disorders

◎藪謙一郎, 伊福部達(東大)

- ◆筆者らは、指やペンで F1-F2 操作盤面上をなぞることで、リアルタイムに楽器のように音声を生成できる音声生成器を提案し、開発・改良を行っている。
- ◆音声・言語機能の障害は、発声機能、構音機能、その制御機能のいずれの障害でも起こり、残されている器官を活かせば、それだけ自由度が高まる。
- ◆本研究では、構音障害を対象とする、本人の声を活かしたウェアラブルな構音障害の支援システムの開発を目指して、試作機を開発を行った。その際、少ない計算量で声質を簡易的に再現する手法を考案し、実装した。

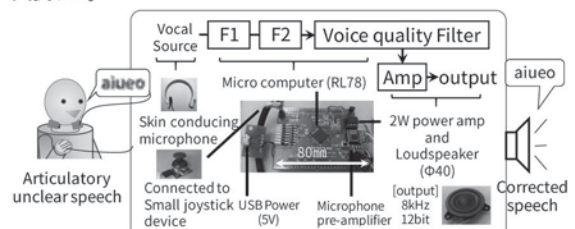


Fig.1: A functional prototype of a voice correction system

### 1-5-15

#### 1-5-15 防災放送における 文節単位の強調の違いによる効果的な伝達 Effective communication by emphasis on each phrase of announcements for disaster prevention.

○高野佐代子 土田義郎(金沢工大)

- ◆防災放送トレーニングシステム構築の一環として、一般人による自然な防災放送「ただちに指定の小学校に避難してください」の「避難して」を強調の効果を確認するため、「大きさ・高さ・速さ・間」を組み合わせた変換合成を行い、聴覚印象の変化について調べた。またプロおよび素人の自然な音声および「逃げる気になる」評価が高い音声の基本周数および速さの分析を行い、総合的な考察を行った。
- ◆一般人が防災放送を行う際、全体を「高く」、また強調部位を「高く」することが重要である。また速くすることにより、「危機感」の評価が上がるが、「重要である」などの評価が下がる傾向がみられた。

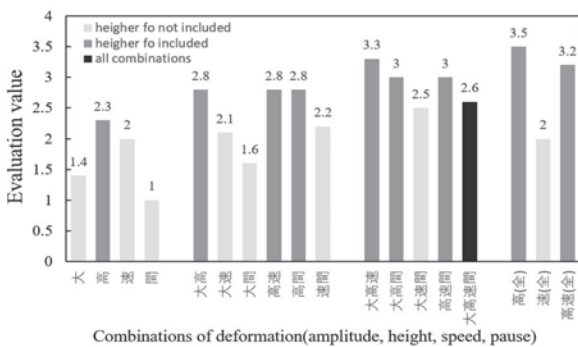


Figure 1 Evaluation results of "desire to escape"

### 1-6-1

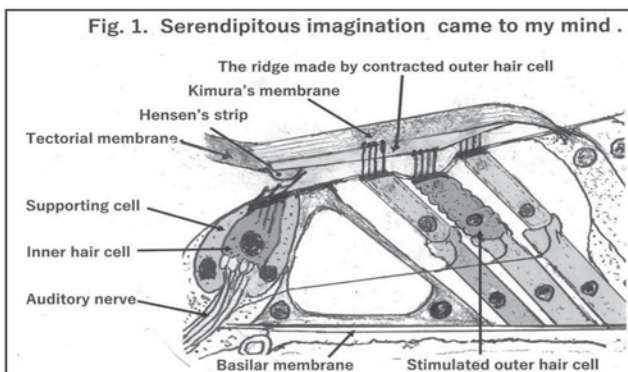
#### 1-6-1 外有毛細胞共振説の聴覚電気生理的説明 Electro-physiological explanation of outer hair cell resonant theory

○森満 保 (宮崎大学 名誉教授)

本報告では、哺乳類蝸牛の綿密な周波数分析能の新機構として、新たに提唱した外有毛細胞共振説を聴覚電気生理学的に検討する。

共振説は、外毛細胞の固有振動数と刺激音の周波数が一致すると、共振が起こり、プレスティン性に短縮し、感覚毛が接着している部位の蓋膜・ヘンゼン条が引き下げられて内毛細胞感覚毛と接触するという新説である。Fig. 1,

電気生理学的に言い換えると、+80mVの内リンパ電位 (EP) をおびたヘンゼン条から、-60mV に帯電している内毛細胞へ電流が生じる。蝸牛マイクロ本電位 (CM) の発生である。ラセン神経節ニューロンには神経活動電位 (AP) が流れる事になる。



### 1-5-16

#### 1-5-16 他人の「マイボイス」を作る How to record other's voices for use in "My Voice".

○竹内京子(國學院大, 日本福祉教育専門学校)

- ◆「マイボイス」という AAC がある。HeartyLadder という文字入力支援ソフトのために「自分の音声」でデータを作成するソフトである。本間武蔵氏(都立神経病院OT)と吉村隆樹氏(パソボラ)を中心に全国の患者家族やサポーターの方々に支えられているフリーソフトである。
- ◆前回の発表(2017年度実施分を発表)で ST 養成校の音響学の授業で音声の音響分析の応用として自分自身の「マイボイス」作成実習を行ったことを報告した。
- ◆その後、2018年度より、実際の臨床を念頭に置き、自分ではなく他人の「マイボイス」を作ることにした。また、完成後の評価としてマイボイス入門コンクールを行った。さらに2019年度は、録音する際に気をつけることを事前に話し合ってもらった。本発表では、これら2年間の報告を行う。
- ◆音響学の授業の合間に、効率よく「マイボイス」実習をできるようにするには工夫が必要である。また、「マイボイス」は実際にどのように使えるかという点についても未知のものである。今後も今回の録音方法についても、学生とともに様々な可能性を考えていきたい。
- ◆「マイボイス」は、ST 養成校の音響学の授業を実際の臨床に具体的に近づけてくれる存在のひとつである。また、音声コミュニケーションの存在意義をあらためて考えるきっかけを与えてくれるものでもある。その手助けができるよう、今後もより活用していきたい。

### 1-6-2

#### 有毛細胞共鳴説

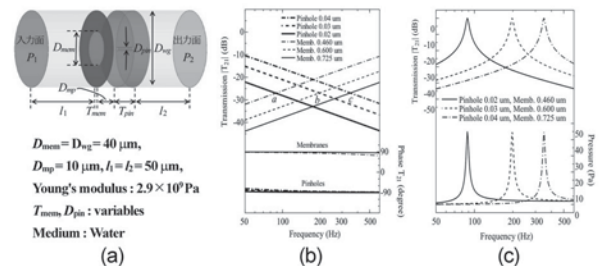
～物理寸法に依存しない音響共鳴の設計～

Resonance theory on outer hair cells

～Design of physical-dimension-free acoustic resonance～

☆玉木愛莉, 洪文甲, 北村敏明, 堀井康史(関西大学)

我々は、音を聴く際に内耳に存在する外有毛細胞が、物理寸法に依存しない音響共鳴を起こすと考え、外有毛細胞共鳴説を唱えてきた。この説は、細胞のルーレット孔がコイル L、細胞膜がコンデンサ C として動作することにより、音刺激がルーレット孔から細胞膜へと伝わる過程で LC 直列共鳴を起こすというものである。しかし、細胞の寸法がミクロン寸法であることから、常識的には波長に基づく共鳴を起こすことは考えられない。本稿では、集中定数回路の共鳴原理に基づき、微細な構造で鋭い音響共鳴を得るための設計法を提案する。



- (a) A resonant model including a pinhole and a membrane.
- (b) Transmission characteristics of pinhole models and membrane models (upper: amplitude, lower: phase).
- (c) Transmission characteristics of resonant models (upper: transmission, lower: inner pressure).





### 1-6-7

**1-6-7 時間的側面から見た音声知覚における感覚運動情報処理の脳内メカニズム: 脳磁図と機能的共鳴画像法を用いた検討**  
 Temporal characteristics of sensorimotor processing in speech perception: Multimodal integration of magnetoencephalography and functional magnetic resonance imaging

◎田村俊介(九大シス情), 廣永成人, 光藤崇子(九大医), 廣瀬信之, 森周司(九大シス情), 帆秋伸彦(帆秋病院), 中村一太, 鬼塚俊明, 平野羊嗣(九大医)

- ◆音声知覚時に起こる音声生成に関する脳領域の活動は、感覚情報から音声信号を作り出す原因となった調音運動を逆推定する処理を表すと考えられている。
- ◆本研究では、調音の時間特徴 (VOT) に特徴付けられる音韻対 (有声音/d/と無声音/t/) のカテゴリー知覚を行っている際の音声生成に関する脳領域の活動とその活動と機能的結合を持つ脳内ネットワークを脳磁図と機能的磁気共鳴画像法を用いて調べた。
- ◆有声音と無声音のカテゴリー知覚精度が高い参加者程、音声生成に関する脳領域の内、声帯振動の開始や準備の役割を担い、VOTの制御に密接に係ると考えられる subcentral area の活動が強く、頭頂弁蓋(OP4)を介した聴覚皮質との機能的結合 (100-300 ms) も強かった。

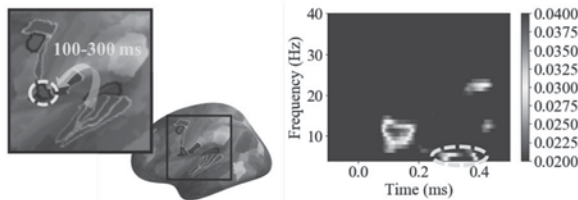


Fig.1: Summary of the present study.

### 1-8-1

#### 1-8-1 流体音計算時に生じる大量の音源に対するFMBEMに基づく高速処理手法

Fast calculation method based on FMBEM for many sound sources generated in flow induced noise analyses

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

- ◆Lighthill方程式に基づく流体音解析を境界要素法 (BEM) で実施する際は、大量の音源を設定するため、各音源の寄与を全境界要素節点で評価する処理 (音源寄与計算) に要する計算時間が問題となる。
- ◆高速多重境界要素法 (FMBEM) で用いられている高速多重極法 (FMM) を音源寄与計算に適用し、処理の効率化を試みた。
- ◆音源寄与計算を従来法で実施した場合の、境界面上における音圧レベル分布をFig. 1に示す。この結果はFMMを用いた場合の結果と良好に一致しており、FMMによる音源寄与計算が正しく行われていることを確認した。
- ◆従来法による計算時間は10,000秒を要するが、FMMの適用により多くの周波数で300秒以下となった (Fig. 2)。
- ◆FMMの適用により追加で必要となるメモリ量は最大2GBであり、近年のCAE実施環境として大きくない。

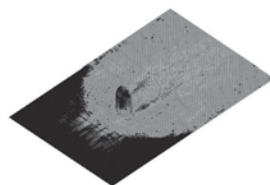


Fig. 1 Sound pressure level distribution at 200 Hz.

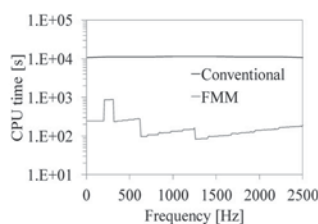


Fig. 2 Computational cost.

### 1-6-8

#### 1-6-8 変形聴覚フィードバックシステムを用いた実時間英語母音発音変換の検討

A study on real time English vowel pronunciation conversion using transformed auditory feedback system

◎廣谷定男, 持田岳美(NTT CS 研)

- ◆非母語発音の訓練においては、学習者がネイティブの発音を聞き、それを真似て発音することが一般的である。
- ◆しかし、我々がこれまでに開発してきた変形聴覚フィードバックシステムを用いれば、学習者が正しい発音に変換された自分の声を聞きながら非母語の発音を行うことができるため、非母語発音の改善が期待される。
- ◆本研究では、日本語母語話者を被験者とし、英単語 "had" および日本語モーラ「は」の母音発音を英語母音[a:]の発音に変換した音声を聴覚フィードバックし、適応させた場合の発音の変化を調べた。
- ◆実験の結果、適応により、「had」では発音の変化が観測され、「は」では観測されなかった。また、適応後に発音に変化し続けることが明らかになった。
- ◆このことから、本手法が非母語発音の訓練に有効な可能性がある。

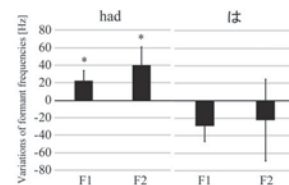


Fig.1: Variations of formant frequencies for "had" and 「は」.

### 1-8-2

#### 1-8-2 Guided-spatio-temporal filterを用いた光学的可視化音場と流れの分離

Extracting sound field from air-flow captured by optical method using guided-spatio-temporal filter.

◎谷川理佐子, 矢田部浩平, 及川靖広 (早大理工)

- ◆時空間フィルタによる流れと音の分離
  - 流れ内部の音場の観測に有効
  - 物体近傍・画像境界付近の音場の抽出が困難 (Fig. 1)
- ◆ガイド付きフィルタを導入した Guided-spatio-temporal filter
  - ガイド付きフィルタ: 画像のエッジを保持したまま平滑化フィルタリング可能
  - 空間フィルタとしてガイド付きフィルタを利用
    - ◇ カットオフ周波数は音の周波数に応じて変更可能
    - ◇ ガイド画像は物体を隠すバイナリマスク画像を使用
- ◆数値シミュレーションと実測データへの適用 (Fig. 2) により物体近傍を含めた音場の抽出が可能であることを確認

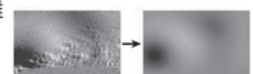


Fig. 1: Original and spatio-temporal-filtered images.

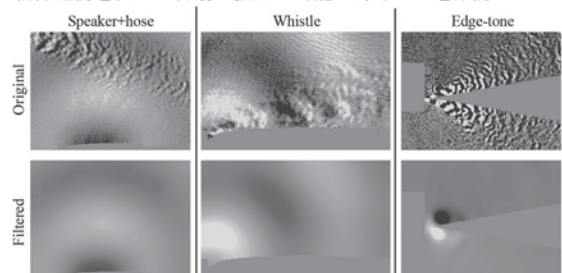


Fig. 2: Original and guided-spatio-temporal-filtered images captured by optical method. Upper images are the original and lower images are the filtered images.

### 1-8-3

#### 1-8-3 高速気流が音の伝搬特性に与える影響 - 簡易風洞を用いた実験的検討 -

Effect on the sound propagation characteristics caused by high speed airflow  
- The study with the simple wind tunnel experiment -

星加 慧, 岩永 景一郎, 土肥 哲也(小林理研)

- ◆ 風速 0~40 m/s の気流によって、音の振幅と伝搬時間がどのように変化するかを調べた。
- ◆ 気流中の音波の振る舞いを明らかにすることで、航空機や新幹線等の移動音源の騒音予測や音源位置同定の精度向上に役立つ可能性がある。
- ◆ 気流の上流側で音圧の増大、下流側で減少する傾向が見られた。

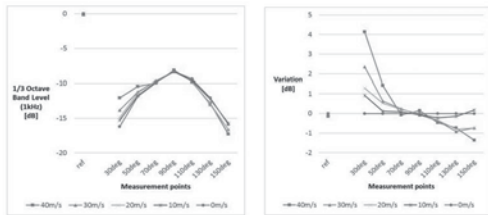


Fig.1 Third Octave Band Level (1000 Hz).

- ◆ 伝搬時間は上流側で遅れ、下流側で早まる傾向が見られた。

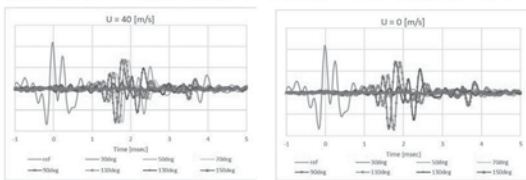


Fig.2 Impulse Response

### 1-8-5

#### 1-8-5 高高度を飛行する航空機からの騒音の伝搬 - 気象条件とレベル変動の関係 -

Sound propagation from aircraft flying at high altitude  
- relationship between meteorological conditions and level fluctuations

○ 牧野康一(小林理研)

- ◆ 高高度 (5~9.5 km) から伝搬する航空機騒音のレベルは航空機の移動に伴った1分程度の長い時間の変動と、短い周期の変動が見られる。
- ◆ ゆっくりとした変動を  $L_{AS}$  の10秒間移動平均で表し、 $L_{AS}$  との差分で短い周期の変動を表した。
- ◆ 短い周期の変動が大きな場合(下図)、高層気象観測データに風速の鉛直分布の大きな変化や気温の逆転層などが見られた。

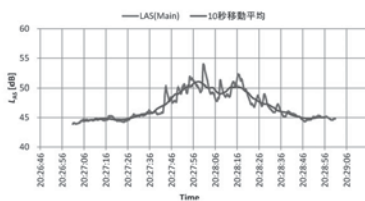


Fig. A 特性レベル波形 ( $L_{AS}$ ) と 10 秒間移動平均波形 (B773 真上 東→西 高度 5600 m, 快晴, 地上風 4m/s, 上空風 11m/s, 逆転層)

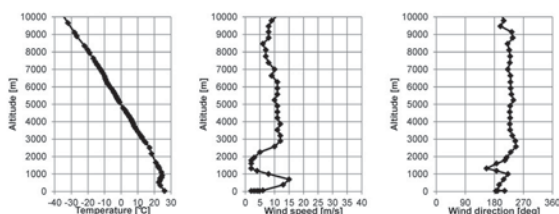


Fig. 測定データに対応する高層気象観測結果

### 1-8-4

#### 1-8-4 成田空港における 大型マイクロホンアレイを用いた ジェット旅客機の音源分布測定を試み

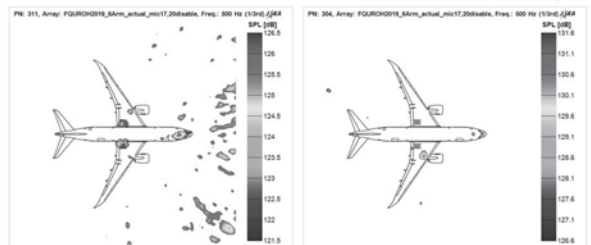
A noise source distribution measurement of jet airliners at Narita airport by using large aperture phased-array microphone

○ 小林知尋、横田考俊、牧野康一(小林理研)、  
△ 加藤貴幸、△ 下田啓司 (ISE)、△ 高石武久 (JAXA)

- ▶ マイクロホンアレイで地上から航空機の音源分布を測定し、その結果から詳細な騒音予測のための音源モデルを作成する
- ▶ 様々な機種別の音源データを収集するための技術検討として、成田空港において旅客機を対象とした予備計測を実施した(Fig. 1)
- ▶ 様々な技術的制約があったが、搭載エンジンの型式違いによる騒音発生特性の違いが分かる結果が得られた(Fig. 2)



Fig.1 Noise source distribution measurement at Narita Airport



(a) B787-800, GEnx-1B70 engine (b) B787-800, Trent 1000 engine  
Fig. 2 Examples of acoustic maps (500 Hz in 1/3 octave band)

### 1-8-6

#### 1-8-6 異なる地表面性状における側方減衰の周波数特性比較

Comparison of frequency characteristics of Lateral Attenuation in different ground surfaces

○ 川瀬康彰、花香和之(成田空港振興協会)、篠原直明(空港支援機構)、  
牧野康一(小林理研)、山元一平(防衛施設協会)

- ◆ 空港側方の地表面性状が田んぼ、畑地、市街地と異なる3空港で測定した離陸滑走音の調査結果に基づき、地表面性状の違いに因る側方減衰量やそのばらつき及びベクトル風速依存性の差異について周波数帯ごとに比較した。
- ◆ 空港間で見られた差異について、ばらつきについては地表面性状の同一性、ベクトル風速依存性については地表面の粗度の違いが原因として考えられた。
- ◆ 側方減衰量の大きさについては、伝搬距離や周波数帯によって空港間で大小関係が異なる結果となった。

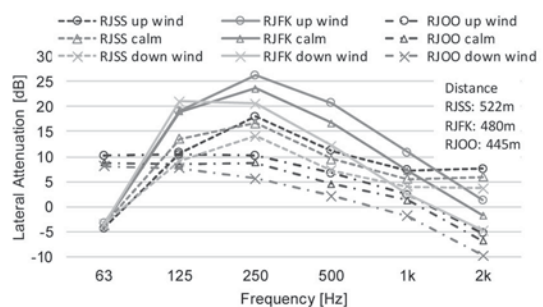


Fig.1 Comparison of frequency characteristics of Lateral Attenuation on equal propagation distance

### 1-8-7

#### 1-8-7 屋外音響伝搬に及ぼす地表面と気象の影響に関するフィールド実験

—インパルス応答自動計測システムを用いた長期間データ収集—  
Field experiment on influences of ground and meteorology on outdoor sound propagation

-Acquisition of excess attenuation data in long-term by using the automatic impulse response measurement system -

○横田考俊, 牧野康一(小林理研), 森淳一, 山元一平(防衛協会),  
△飯泉元氣, △堤拓哉(北総研)

- ◆屋外音響伝搬に及ぼす地表面および気象の影響を把握することを目的とし, 2018年10月より, 屋外の定点間(伝搬距離300m)において長期間に渡り継続的にインパルス応答の計測を開始した。
- ◆商用電源を要さないインパルス応答自動計測システムを用いて実験を行い, インターネットを介して遠隔地においてデータを収集した。
- ◆1日3回の定時(9時, 12時, 17時)にインパルス応答計測を行うとともに, 風向風速, 気温, 湿度, 大気圧および積雪深を把握した。
- ◆超過減衰(E. A.)を算出してベクトル風速との関係について検討するとともに, 積雪深と地面反射音レベルの関係について検討した。

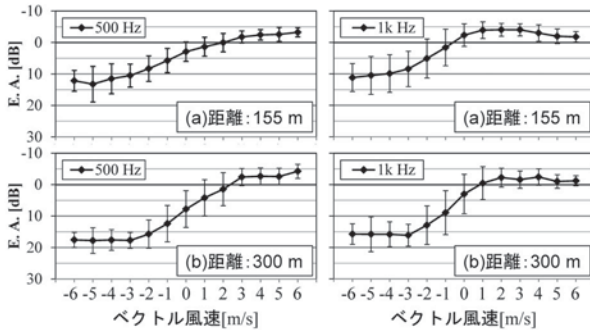


Fig.1: Relationship between vector wind speed and excess attenuation.

### 1-8-9

#### 1-8-9 防音壁からの透過音の影響を考慮した騒音評価手法

Acoustic FEM Analysis Considering Difference in Sound Insulation Performance of Noise Barrier

◎佐藤大悟(鉄道総研/九大・芸工), 尾本章(九大・芸工)

- ◆数値シミュレーション内で防音壁の遮音性能の違いを反映することを検討した。本研究では, 有限要素法を活用して解析領域中に通常の音の伝搬空間である空気とは異なる媒質を設け, 媒質の複素波数によって防音壁等に用いられる遮音板の透過損失に基づく音の減衰効果を付与した。
- ◆減衰媒質の複素波数を所望の音圧レベル低減量との関係に基づいて決定した結果, 平面波伝搬モデルによって減衰媒質前後での音圧レベル減少量が付与した効果と一致することを確認した。
- ◆直壁モデルや鉄道高架橋を対象に, 同手法を用いた解析を行った結果, 音圧レベル分布は透過音の影響と考えられる変化を示し, 透過音を考慮した音響解析手法として有用と考えられた。

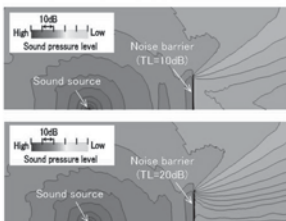


Fig.1: Comparison of SPL distribution based on the difference in sound insulation performance

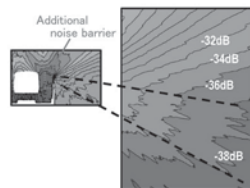


Fig.2: Influence of transmitted sound from additional noise barrier

### 1-8-8

#### 1-8-8 ANSI/ASA S1.18-2010 法およびアンサンブル平均法による地表面音響特性の実測比較

Comparison of measurement results of ground properties by the ANSI/ASA S1.18-2010 and the ensemble averaging methods

☆岩本拓也(新潟大・院), 大嶋拓也(新潟大・工), 黒坂優美(新潟大・院)

- ◆地表面音響特性を現場で測定可能な手法は国内外で多数開発されているが, 実測による手法間の相対的な評価は不十分である。
- ◆そこで, 国外で知られるANSI/ASA S1.18-2010法(ANSI法)および, 国内で多用されるアンサンブル平均法(EA法)を用い, 草地および土の地表面を, 同日に2回および別日に2回(Fig.1)測定する。
- ◆ANSI法Step1では, 同一のテンプレートに適合する限り同一の結果が得られる(Fig.1(a)-(c))。同法Step2では, 安定した結果が得られない((d)-(f))。EA法は音響特性の小さな変化を捉えるが, 同時測定した土の含水率から想定される変化と整合しないことがある((g)-(i))。

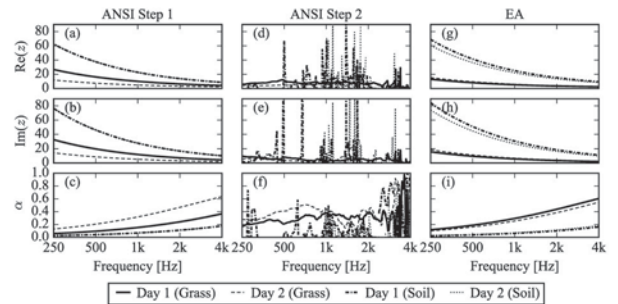


Fig.1 Impedance and absorption coefficient of grassy and soil grounds obtained in different two days.

### 1-8-10

#### 1-8-10 福岡市における交差点付近での警笛の使用状況の調査

Survey on Vehicle Horn Use at Intersections in Fukuoka

☆小林和歩, 常川翔貴, 橋本和真, 福田環,

高田正幸(九大・芸工), 金基弘(駿河台大), △大枝良直(九大・工)

- ◆本研究は自動車の警笛が使用される要因を明らかにし, 警笛の使用を必要最小限に抑えるための方策を提案することを最終的な目標としている。本稿では, 交通量や信号周期といった交通環境上の特徴と警笛使用の関係性を調査, 検討した。これまで台湾で行ってきた調査と比較するため, 本調査は, 福岡市の三つの交差点を対象に行った。
- ◆総交通量と警笛の発生に対応を検討した結果, 一部の交差点では対応が見られた。
- ◆車種ごとの交通量と警笛の発生に対応を検討した結果, 普通車の交通量と警笛の間に関連がある可能性が示された。
- ◆信号周期と警笛の発生を関連を検討した結果, 各交差点で信号フェーズが変わるタイミングでの警笛の発生が著明に増加することが確認できた(Fig.1)。

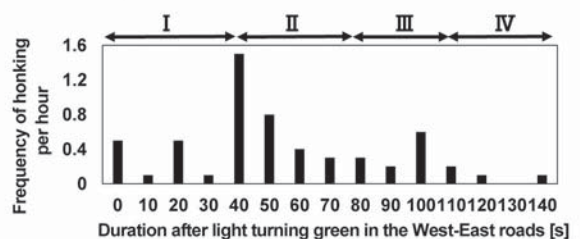


Fig.1: Temporal variations of frequency of honking after traffic light turning green

### 1-8-11

#### 1-8-11 深層学習を用いた違法騒音車両の判定

Determination method of illegal noise vehicle from pass-by noise by using deep learning  
○宝渦寛之(交通研)

- ◆違法騒音車両の街頭検査に用いられる試験法は、車両を停止させた状態で行われるため、走行騒音とは相関がない。そこで、本研究では、深層学習を用いて、走行騒音から街頭検査で違法と判定される車両を判定可能を検討した。
- ◆深層学習には AlexNet を使用し、“違法車両”、“適法車両”、車両以外の騒音である“その他”の3つのカテゴリに分類を行う教師あり学習を行った。
- ◆“違法車両”、“適法車両”の学習データを取得するため、テストコース上で、公道走行不可のマフラーを装着した状態で、騒音測定を行った。
- ◆取得した騒音データは、一定時間幅ごとに、横軸を時間、縦軸を周波数とした画像に変換した。
- ◆検討の結果、速度、加速度等が任意の状態の車両に対して、約95%程度の精度を有する学習モデルを作成することが出来た。

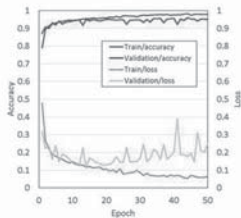


Fig.1: Accuracy and loss of calculated model by deep learning

### 1-8-13

#### 1-8-13 道路の舗装種別が走行騒音に与える影響について

Effects of road pavement type on vehicle traveling noise

○岩永則城((株)計算力学研究センター)、明石克人(飛揚航空機製造開発(株))

- ◆波動音響解を用いて、車両走行騒音における道路舗装タイプの影響を調べた。
- ◆3種の舗装タイプに対し、車両がマイクに近づくにつれて騒音レベルが上がり、通過後は下がるという傾向は、波動解と幾何音響解に共通であった
- ◆密粒舗装のユニットパターンは、高性能舗装 II 型に近いが、排水性舗装とは異なっていた。

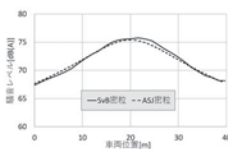


Fig. 2 密粒舗装ユニットパターン

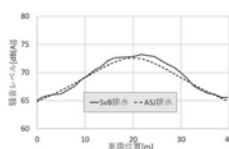


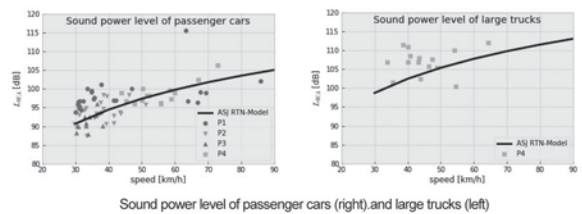
Fig. 3 排水性舗装ユニットパターン

### 1-8-12

#### Measurement of sound power levels of free running vehicles in São Paulo

☆ Marjorie Takai (Grad. Sch of Eng., UTokyo), Hyojin Lee (IIS, UTokyo), Shinichi Sakamoto (IIS, UTokyo)

Developing countries, such as Brazil, have been enforcing the use of noise maps following the trends of EU, including their methods and models as framework for investigation. This paper presents a screening of the sound power levels of free running vehicles in São Paulo using the novel system, composed by a SLM and video camera, and the Japanese ASJ-RTN methodology and category classification. The sound power levels measured are higher than the values given by the ASJ RTN-Model for passenger cars, around 1.8 dB, and 4.5 dB in for large trucks. Specially the later, is expected due to larger size of vehicles itself when comparing with Japan, even not considering vehicles with more than 3 axles. However, more measurements are necessary for a more complete assessment of the sound power levels emitted in the different scenarios in Brazil.



Sound power level of passenger cars (right) and large trucks (left)

### 1-8-14

#### 1-8-14 遮音壁が併設された桁構造高架道路裏面反射音の伝搬予測 —3次元波動数値解析とASJ RTN-Modelによる結果の比較—

Prediction of road traffic noise reflected from the bottom of an elevated road structure with I-type girders and diffracted from a roadside barrier: comparison of results by 3-D wave-based numerical analysis and ASJ RTN-Model

☆大谷知光, 関根秀久, 安田洋介(神奈川大・工), 福島昭則(ニューズ環境設計)

- ◆高架道路下に地表面がある場合や遮音壁が設置された場合を想定し、ASJ RTN-Modelの散乱反射法と3次元波動数値解析手法を用いてユニットパターンやその積分値、音圧レベル分布を算出し比較することで、ASJ RTN-Modelの予測精度の検証を行った。
- ◆自由空間内 (Case A) と地表面のある半自由空間内 (Case B) では、ASJ RTN-Modelと3次元波動数値解析の結果が良好に一致した。
- ◆地表面上に遮音壁が併設された場合 (Case C)、回折減衰の大きい領域で両手法の結果に乖離が見られた。

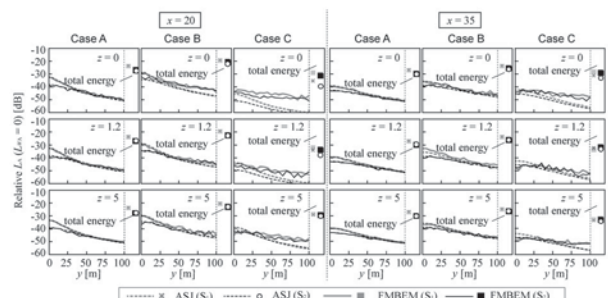


Fig. 1 Unit patterns of  $L_A$  and the values of energy integration.

## 1-9-1

1-9-1 聴能形成課題における  
客観的難易度と主観的難易度の関係Relation of Objective Difficulty and Subjective Difficulty  
in Technical Ear Training Task

○丸井淳史, 亀川徹(東京藝大)

- ◆聴能形成(音の判別・識別能力向上訓練)は、音に対する鋭い感性を養うために、音の物理的な変化と感覚的な変化とを結びつけて理解するための訓練方法である。
- ◆訓練に用いる音源の選択は教師の暗黙知に委ねられていることが多く、学習者にとっては、どのような音源が聴能形成において効果的であるかの判断をすることは難しい。
- ◆前報では音源による**主観的難易度**に寄与が高かったのは、音源のパワースペクトルの回帰直線からの残差の二乗平均平方根(RMSres)であった。
- ◆そこでRMSresを調整した音源を用意し、**客観的難易度**にどのような影響があるかを調査した。
- ◆その結果、RMSresを調整した音源では調整前のものと比べて正答率が低くなった。
- ◆主観的難易度が低くなるような調整が客観的難易度を高めることは意外であるが、プロのエンジニアの耳によって完成された音を改変したことが聴能形成での正答率に悪影響を与えたと考えられる。

## 1-9-2

1-9-2 複数名が声を合わせる場面での  
音の可視化装置の教育への活用

Application of sound visualization device to sound education in chorus

○齊藤忠彦(信州大・教), 財満健史, 大脇雅直(熊谷組技研),  
田島達也(信州大・教)

音楽教育における歌唱指導の場面で、従来の聴覚情報に加えて音の可視化装置を用いた視覚情報の活用を検討している。本報では、音の可視化装置による視覚情報を活用する方策について、4名の被験者が声を合わせる練習を行い、その初期段階と練習を重ねた後を比較した。被験者らは、練習を重ねた後の方が声は合っていたと自己評価している。可視化装置による結果は、練習量に関わらずピッチ(音の高さを示す円の色)はおおよそ合っていた。練習を重ねた後の方が、声量(音量を示す円の大きさ)のばらつきは少なかった。また、音源方向を示す円が被験者の間で揺れ動くように見えた。これは声量やピッチのばらつきが少なくなったためと考えられる。視覚情報の活用により、声量のバランスやピッチを客観的に確認することができた。さらに声が合っている状態においては、音源方向を示す円が被験者間で揺れ動くように見えることから、歌唱指導の場面で効果的に活用することができる可能性を示した。

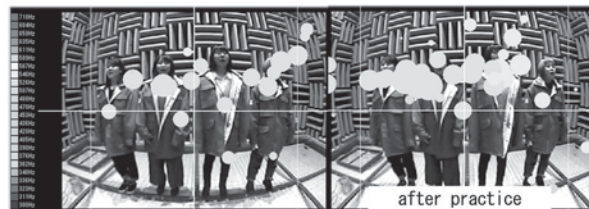


Fig.1: Results of visualization (Oct. band 500Hz and under, C3)

## 1-9-3

1-9-3 バーコードリーダーによる母音スペクトログラムの  
パターン・プレイバック

Pattern playback of vowels by barcode reader

○安啓一(筑波技術大), 荒井隆行(上智大)

- ◆スペクトログラムを音声に再合成するパターン・プレイバックとしてバーコードリーダーのアナログ出力を使用することを試みた。
- ◆レーザースキャンの1周期分の濃淡情報を100 Hz-8000 Hzの80帯域に分割し、対応する周波数のcos波を重ね合わせ音声を再合成した。
- ◆/a/, /i/, /u/, /e/, /o/の5母音のスペクトログラムから音声の再合成を行った。
- ◆今後はバーコードリーダーで読み取った情報をリアルタイムに音声に変換することを目指す。

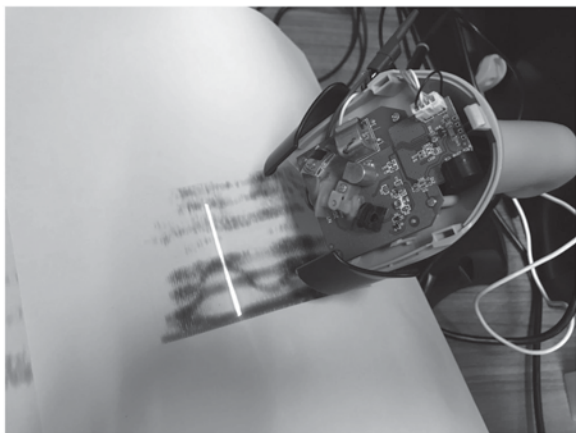


Fig.1: バーコードリーダーによるスペクトログラムの読み取り

## 1-9-4

## 1-9-4 おやつを用いた音育の可能性

Possibilities of sound education using snacks for children

○木谷俊介(JAIST), 宮田大(Konel)

- ◆ 保育施設で行ったワークショップについて紹介する。
- ◆ おやつを噛んだときの音を擬音化することをコンセプトとした「オノマトペのおやつたち」を用いて、おやつを食べながら音を意識することを試みた。
- ◆ ワークショップでは、1) 子どもたちが能動的に静かになる状況を作ること、2) 音を感じることができること、3) 音を発することができることの3つを達成目標とした。
- ◆ ワークショップは、「おやつを食べてもらい、そのときどんな音が自分の口から聞こえるかを答えてもらう」というものであった。
- ◆ 結果は、1) 静かにすることを促すことなく静かになった、2) 耳に手を当てるなど音に集中していた、3) 各自の言葉でオノマトペ化できていた。
- ◆ おやつを用いることで、保育園に通う年代の子どもたちに楽しみながら音を意識し、音に耳を傾けてもらう試みができたと考えられる。

1-9-5

1-9-5 作って学ぶ音響教材「スマートフォン拡声器」を活用した音の工作教室および高専出前授業の実施報告

Practical reports of craft workshop and delivery lectures of technical college utilizing the acoustic teaching aid "smartphone stand having acoustic horn"

○高橋 義典, 伊藤 沙織, 小林 奏人, 齊藤 章太, 真志取 秀人 (都立産業技術高専)

- ◆筆者らが音の伝達に関する実験教材として提案してきた、音響ホーンの工作教材「スマートフォン拡声器」を、1時間程度で組み立て可能なペーパークラフトの工作教材に改良した。
- ◆提案する教材を利用して、日本音響学会音響教育委員会主催「音の工作教室」と、東京都立産業技術高等専門学校主催「出前授業」を2回、合計3回の音響教育イベント実施した。
- ◆本教材の組み立ては、小学生にとっては難しく感じられる場合があるものの、工作は十分に楽しめた様子が確認できた。従って、工作教材には程よい難易度が必要であると考えられる。

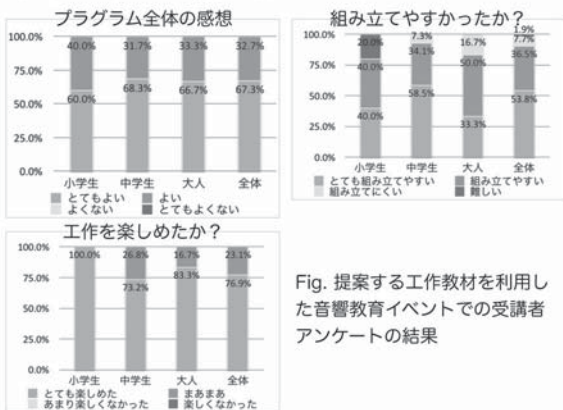


Fig. 提案する工作教材を利用した音響教育イベントでの受講者アンケートの結果

1-9-6

1-9-6 この音何デシベル?-大学生の音の大きさに対する理解度をまずは測る試みその2-

What is this sound in dB?

We tried to determine student's comprehension of sound levels Part2.

○上田麻理, 西口磯春 (神奈川工科大学)

<あらまし>

本稿では、大学生の音の大きさとその単位のデシベル(dB)に対する理解度を測る試みとして、自作した音の大きさチャート・キットを用いて理解の具合を測ってみた。今回は音楽大学生を対象として実施した。

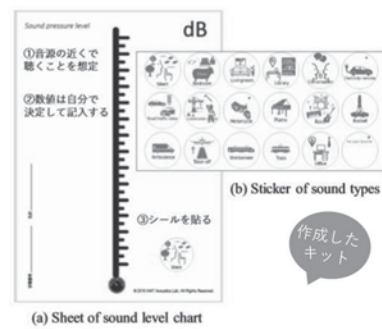


Fig. Detail of the sound level chart kit. (a) sheet and (b) stickers of sound types (18 types).

1-10-1

1-10-1 パラレルフィルタを用いたピアノダンパーペダルのモデル化

Modeling a piano damper pedal using parallel filters

☆小林憲治, 矢田部浩平, 及川靖広(早大理工), 宮城雄介, 大石耕史(株)コルグ

- ◆研究目的
  - > ダンパーペダルによる共鳴を少ない計算量でモデル化する
- ◆提案手法
  - > パラレルフィルタによって全ての共鳴弦をモデル化
  - > ピアノ収録音からモデル化に適した極を推定・選出
- ◆数値実験
  - > パラレルフィルタモデルによりダンパーペダルの共鳴音を生成
  - > 従来手法 (FZ-ARMAモデル)と比較し、表現力の高いモデルが設計できていることを確認

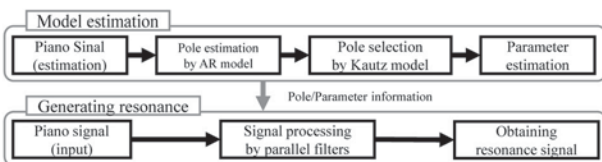


Fig. 1: Diagram of proposed method

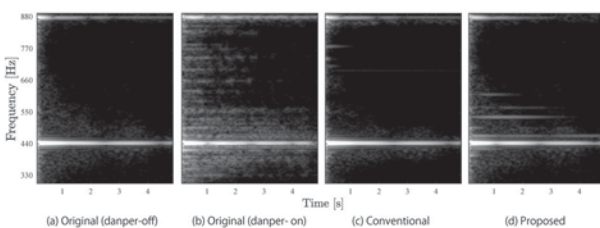


Fig. 2: Magnitude spectrogram of each signals

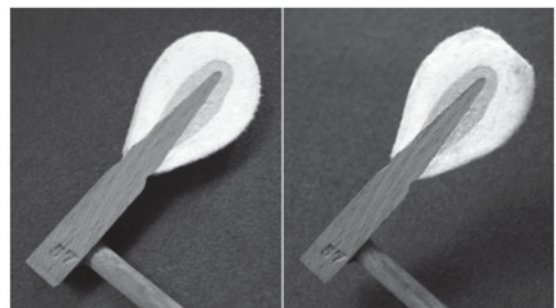
1-10-2

1-10-2 ハンマーの経年劣化を考慮したグランドピアノの物理モデリング音源

Physical Modeling Sound Synthesis of Grand Piano Considering Hammer Aging.

☆関口周, 鮫島俊哉 (九大・芸工)

グランドピアノのハンマーは、演奏を繰り返すと、フェルト部が金属弦に当たることによりすり減り、形状が変化してゆく。この形状の変化をモデルに取り入れることにより、経年劣化をモデル化することを試みた。結果は、すり減って形が平らになるほど高域が下がるものとなった。その傾向が理論式と一致することが確認された。



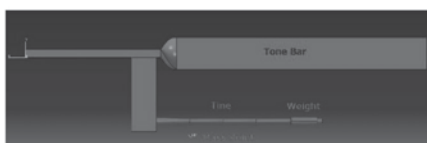
### 1-10-3

#### 1-10-3 電気ピアノ振動体の物理モデルに関する検討

On physical modeling of the vibrating assembly of an electric piano  
○門倉彩人(神工大), 竹淵瑛一(神工大), 西口磯春(神工大)

◆報告者らの研究室では、ローズピアノの物理モデルについて検討し、振動体の簡易モデルと電磁場解析結果を利用した物理音源モデルを作成した。現在、その精度向上を目指した検討を行っている本稿では、前報に引き続き、ローズピアノの振動体について検討した結果を報告する。

◆ローズピアノの物理音源モデルについては、これまでにも検討されているが、ハンマーで打撃されるTineを片持ち梁としてモデル化してあり、Tineともに非対称の音叉を構成するTone Barの振動については考慮されていない。



ローズピアノの発音部

◆本研究では、3次元FEM解析により、Tone Barの影響について検討した。

◆その結果、Tone Barの存在により打撃方向とは異なる方向の振動が励起されるが、その程度は小さいことがわかった。

### 1-10-5

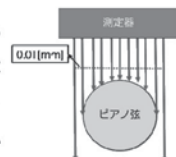
#### 1-10-5 インラインプロファイル測定器によるピアノ弦の空間運動の検討

On the measurement of the two-dimensional vibration of piano strings  
○古知悠太郎(神工大), 林樹裕(鈴木楽器製作所), 西口磯春(神工大)

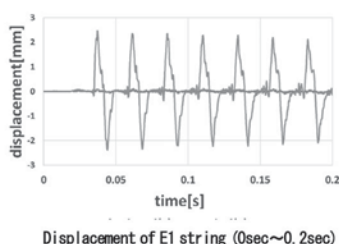
◆ピアノ音は打鍵直後の音量の減少は大きいですが、その後は緩やかに減少していく。この2段減衰の要因として挙げられるのが、弦の空間運動および複数弦の連成である。

◆前者については、駒についたピンの傾きや駒の運動などの影響で水平方向の動きが加わると考えられている。この時、上下方向の運動が2段減衰の前半の急速な減衰に対応し、水平方向の運動の影響が後半の緩やかな減衰に対応する。

◆しかしながら、これらの要因の定量的な理解のためには、実験による弦振動の測定データが不可欠である。本稿では、インラインプロファイル測定器を用いた弦の空間運動の測定法について検討した。



◆これまでに測定した弦は、E1, A1, C2, E2, A2, A3, C4, E4, A4, A5, C6である。サンプリング周波数を16kHzとし、打撃直後約0.5秒の水平・垂直方向の変位を求めた。



### 1-10-4

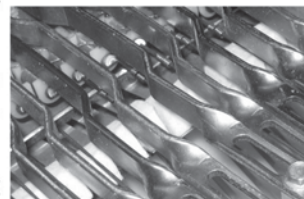
#### 1-10-4 電気楽器用電磁ピックアップの物理モデルに関する検討

On physical modeling of the electromagnetic pickups of an electric piano  
○竹淵瑛一(神工大), 古知悠太郎(神工大), 西口磯春(神工大)

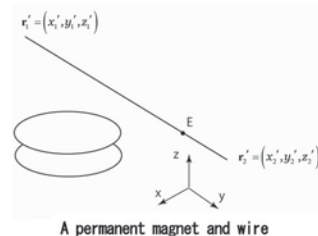
◆本報告は、電気ピアノは1970年代にジャズやロックなどで盛んに用いられたローズピアノの物理モデルを対象としている。

◆前報に引き続き、ローズピアノの電磁ピックアップのモデル化について検討した結果を報告する。

◆電磁ピックアップの先端部には磁極があり、その後方に永久磁石が設置されている。ハンマーにより打撃されたTineの運動によりピックアップコイルに誘導起電力が生じ、電気信号が生じる。



◆前報で行った3次元FEMによる鎖交磁束の計算は詳細な形状をモデル化できる点で優れているが、モデル作成の手間等を考えると、簡易的な方法も魅力的である。ここでは、エレキギターのピックアップの解析法として提案されたHortonらの方法を拡張した方法を検討した。



### 1-10-6

#### 1-10-6 梵鐘の音響特性に及ぼす形状の影響

Influence of the Shape of Bonsho (temple bell) on the sound features  
○中川一人, 塩川博義, 豊谷 純 (日大)

◆本研究では、奈良・薬師寺の梵鐘を対象とし、有限要素法による固有振動解析および1/10モデルの試験体を用いた実験を行い、駒の爪をはじめとした各形状因子が音響特性に与える影響について調べた。

◆有限要素解析結果より、同一形状の振動が、角度がずれて対をなし、この二つの固有モードの僅かな振動数差が梵鐘の音色の大きな特徴であるうなりを作りだしていることが考えられる。しかしながら、装飾的な部位の影響がみられたことから、より正確な解析を行うためには、装飾的な部位を含めたモデリングが必要であると考えられる。

1/10のモデルを用いた実験では、駒の爪と固有振動の関係について調べた。結果、鉛直方向と円周方向を比較すると、円周方向の厚みが与える影響が大であった。

駒の爪の形状は開口部の剛性を上昇させる形状とすることが、効果的に振動数を上昇させると考える。

本研究の対象である薬師寺の梵鐘は奈良時代のものであるため開口部より円周方向に対して厚みが一定であるが、後年には開口部のみが厚い駒の爪が確認できる。これ形状は開口部のみに厚みをもたせることにより剛性を高め、他の部分は軽量に仕上げることもできるため、より効率的に形状であることが考えられる。



### 1-10-7

#### 1-10-7 スティックの多自由度系動特性を含めたシンバルの非線形物理モデリング音源

Nonlinear physical modeling sound synthesis of cymbals involving dynamics of a stick as a multi-degree-of-freedom system

○鮫島俊哉(九大・芸工)

◆シンバルは、Fig. 1に示すような体鳴楽器であり、金属を材料とするシンバル本体と、それを支持するフェルトなどを材料とするワッシャーで構成されている。

◆Fig. 2に示すように、シンバル本体を浅い球殻の非線形振動場として物理モデル化する。ワッシャーは1自由度の機械インピーダンスとして物理モデル化する。Fig. 3に示すように、スティックは1端が回転



Fig. 1 Cymbal, washer and stick.

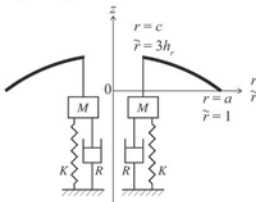


Fig. 2 Cymbal modeled as a spherical shell and cymbal washer modeled as a mechanical impedance.

パネで支持された Euler-Bernoulli はりとして物理モデル化する。スティックとシンバルの相互作用は、はりの衝突振動の解析モデルを適用する。

◆この振動系の時間発展を、差分法と有限要素法によって数値計算する手法を構築する。

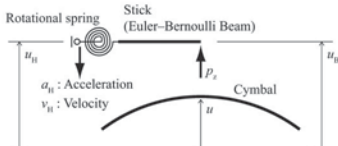


Fig. 3 Stick modeled as an Euler-Bernoulli beam.

### 1-10-9

#### 1-10-9 箏系の材質による押し手(押絃)時の筋活動

On the muscle activities caused when string-pressing of Japanese Koto by difference in material of strings

○安藤珠希(箏曲), 小幡哲史(ヤマハ), 饗庭絵里子(電通大)

◆箏は糸の種類や糸締め状態により、音色はもちろん、弾きやすさも変わることを熟練者は感じている。

◆そこで箏の弾きやすさの指標のひとつとして、箏を演奏する上で重要な動作である、左手で絃を押すことにより13本の絃にはない音高を作る奏法(押し手)に着目した。箏本来の絹糸と化学繊維の糸について、どのような筋活動が行われているのかを計測した。

◆その結果、どの条件においても手首や指を曲げる筋、手首や指を伸ばす筋、肘を伸ばす筋が主として使用されていた。

◆絃の材質の違いに対する有意差を求めたところ、元の音高から半音上げる「弱押し」と一音上げる「強押し」では有意な筋が異なっていたことから、より押絃力が必要な時に対応する筋活動を見ることでも押しやすさの指標のひとつとなる可能性が示唆された。

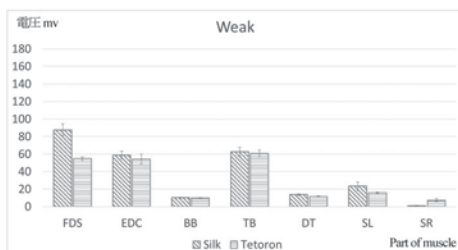


Fig.1: Resultant of muscle activities by material of strings when half tone up.

### 1-10-8

#### 1-10-8 カメラを用いた擦弦楽器における弓圧及び擦弦位置の同時計測

Simultaneous measurement of bow force and bow position in bowed string instruments using a camera

☆猪崎遼大(筑波大院・シス精工)

若槻尚斗, 水谷孝一, 善甫 啓一(筑波大・シス情系)

◆擦弦楽器における弓圧及び擦弦位置の同時計測手法を提案する。

◆弦につけたマーカ―及び弓毛と弦の接点を弦に対して、斜めに設置したカメラで計測を行い、マーカ―の変位を画像処理によって、算出する。その後、擦弦位置をマーカ―の変位との関係から求め、それを利用して弓圧及び摩擦力による擦弦点の変位を算出する。その後、変位を力に変換することで、弓圧及び摩擦力を求める。

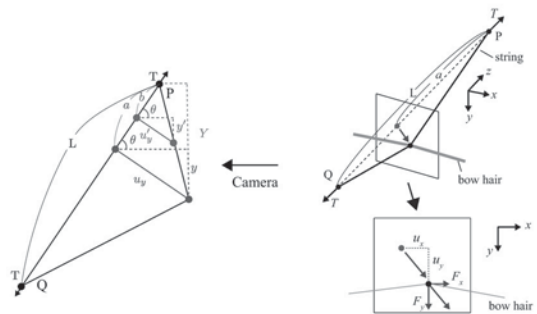


Fig 1 String displacement in the bow force direction Fig 2 String displacement by bow motion

### 1-10-10

#### 1-10-10 異なる鍵盤楽器に対する演奏動作の解析\*

Analysis of performance motion of different keyboard instruments

○三戸勇氣, 川上央, 駒澤大介(日大・芸術), 三浦雅展(国立音楽大学), 松崎桃子, 山口達也, 篠田之孝(日大・理工)

現在まで、我々は感情表出時の楽器演奏動作について検討を行ってきた。今回、3種類の鍵盤楽器の演奏動作の相違について検討を行った。

その鍵盤楽器は、グランドピアノ、キーボード、チェンバロである。

グランドピアノやキーボードは、鍵盤を押す力によって音に強弱をつけたり、音質を変化させることが可能である。しかし、チェンバロは、楽器の特性上、鍵盤を押す力による違いが、音や音質に変化をもたらすことができない。つまり、鍵盤楽器の種類によって演奏動作が変化する可能性がある。そこで、実際に異なる3種類の鍵盤楽器における同一奏者の演奏動作をモーションキャプチャシステムにより計測し、異なる鍵盤楽器が、どのような動きをするのか検討したいと考えた。本研究では感情面に対する3種類の鍵盤楽器の演奏動作をモーションキャプチャシステムで計測し、演奏動作の違いを検討した。今回は、頭部と左右の肩の動きに着目して分析を行った。

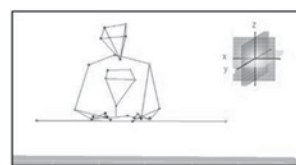


Fig.1 Marker adhesion position

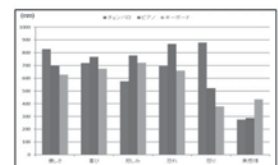


Fig.2 Trajectory length of mean position for four markers of the head

### 1-10-11

#### 1-10-11 f 分の 1 旋律の周波数成分ごとの位相関係

Relation of phase on each frequency component on 1/f melody lines

○三谷尚, △森田卓哉(福岡教育大学)

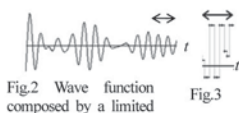
◆倍音の影響のない単旋律の楽曲における音圧関数に関する zero level cross (ZLC) 密度  $h(t)$  に (今後、波形とも呼ぶ) については、多くがその振幅スペクトルについて  $1/f$  型:  $\tilde{A}(f) \propto f^{-\gamma}$  を持つ。併せて、位相情報  $\phi_k$  をも取り込み、フーリエ合成 (フーリエ逆変換)  $\tilde{h}(t) = \sum_k \tilde{A}_k \sin(2\pi f_k t + \phi_k)$  すると、元の波形が再現さ元の波形が再現される:

◆この周波数の和を限られた領域で取り、位相も加味するとどのような波形が得られるかを調べる。  
◆Bach のピアノ小品、Marche による振幅スペクトルは、Fig.1 の様に  $1/f$  型を持つが、その中のサブピーク領域("G", 図中の黒部分)の周波数成分をピックアップし、その領域の振幅と位相を用いて波形を再



Fig.1 Spectrum of Amplitude. (Black region brings wave on Fig.2.)

現した:  $\tilde{h}^G(t) = \sum_k \tilde{A}_k \sin(2\pi f_k t + \phi_k)$ .  
◆その結果、 $\tilde{h}^G(t)$  として Fig.2 の様な波形を得たが、この時間的周期は、この曲 "Marche" の波形に見られる往復的な跳躍 (Fig.3 参照) の周期にほぼ一致。Fig.1 のサブピークが、この音型を表していることがわかった。



◆聖歌 "喜びたえよ" のある周波数域の振幅スペクトルのサブピークを用いて、フーリエ合成をした。これは時間的に周期的であり; 周期はほぼ 1 小節の長さであり、1 小節のフレーズに対応する。

### 1-10-13

#### 1-10-13 構造的音色に基づく新たな空間エフェクト音 —音脈モワレの提案

A new spatial acoustic effect based on structured timbre: a proposal of sound stream moiré

○小坂直敏(東京電機大)

- ◆信号処理技術による近年のエフェクト音合成技術として、モーフィング、サウンドハイブリッドなどが知られている。
- ◆これらの全体の音色は、瞬時の音色の経時変化、組み合わせなどにより構造的に成り立つ (構造的音色)。その結果、一つの音と複合音との中間的に知覚される拡張音脈を形成する。
- ◆以上の観点でエフェクトや現代音楽の特殊奏法を整理。
- ◆新たなエフェクトとして知覚的階層性を持つサウンドコラージュを検討中。これを空間音響エフェクトに適用。音脈モワレ表現と定義。
- ◆音脈モワレ表現の別の具体化: 音声や楽音信号を音色が異なるように 2 つに分解し、これを異なるチャンネルに割り当て再生する方法。
- ◆受聴位置により音色が異なることが示唆される。

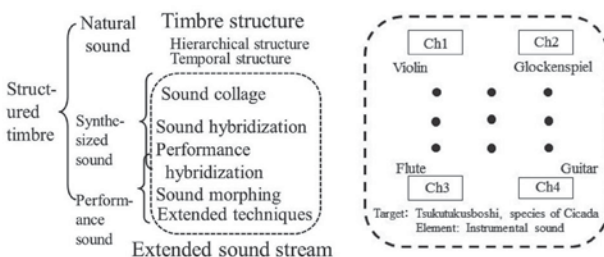


Fig.1 Structured timbres Fig.2 An example of Sound stream moiré where sound collage is applied to spatial acoustics

### 1-10-12

#### 1-10-12 制約極零モデルによるシンセサイザ音源

Waveform synthesis technique using constrained zero-pole model

☆大木大夢, 矢田部浩平, 及川靖広(早大理工), 宮城雄介, 大石耕史(コルグ)

- フィルタ処理を発展させたシンセサイザの新たな音源方式を提案
- ◆フィルタ処理は状態空間モデルで一般化され、自由な取り扱いが可能となる。
- ◆状態空間モデルに非線形な変換を導入することで、新たな合成音を生成できるのではないかと考えた (Fig.1)。
- ◆状態空間モデルを生成するために制約極零モデルにより楽器音の振幅スペクトルを模倣するような IIR フィルタを生成。
- ◆提案手法により音が合成できることを確認した (Fig.2)。

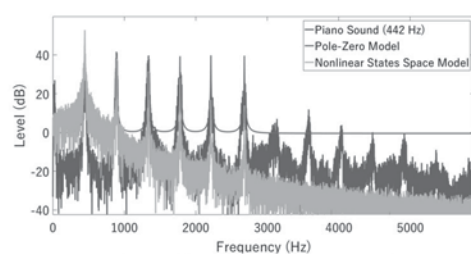
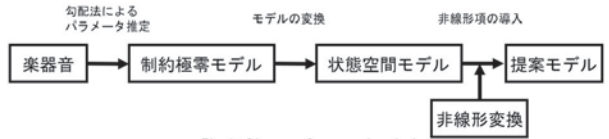


Fig.2: Amplitude spectrum of original sound and synthesized sounds

### 1-10-14

#### 1-10-14 分割音声のステレオ提示時おけるグルーピング知覚特性

Perceptual Grouping Characteristics of Decomposed Speech in Stereophonic Listening

○河野有美, △檜野幸志郎, 小坂直敏(東京電機大)

- ◆本稿では、マルチチャンネル上の新たな音響空間として、音脈モアレ表現の検討を行った。
- ◆ヘッドフォン、ステレオスピーカを用いた心理実験を行い、正弦波モデルによって分解した音声の分離度を評価した。
- ◆男女 2 名の朗読音声を用いた分解音声、混合比率を変えてステレオスピーカから受聴し、5 段階で評価した群化度を Fig.1 に示す。
- ◆混合比率を減じると左右異なるものを聴き、群化度が下がることがわかった。

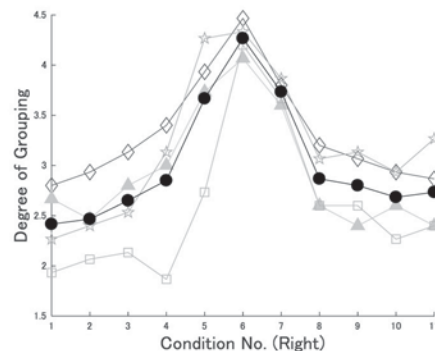


Fig.1: Degree of grouping for decomposed speech in stereophonic listening. (Voices ◇:Female 1, ▲:Female 2, ☆:Male 1, □:Male 2, ●:Ave.)

### 1-10-15

#### 1-10-15 友人との会話におけるBGMの効果

Psychological effects of background music on a conversation with a friend  
○松本じゅん子(長野県看護大)

- ◆目的: 音楽は感情や行動に働きかける効果があるが、実生活でBGMがどのように使われているかは明らかになっていない。日常生活において、自室で友人と会話をする際のBGMの利用と効果について調べた
- ◆方法: 大学生154名を対象に、自分の部屋に友人がいる時にBGMを流すかどうか、流す場合、使用している音楽の種類や使用する理由、効果などについて、質問紙を用いて尋ねた。
- ◆結果と考察: 部屋に友人がいる際にBGMはあまり多くは使用されていないが、使用している場合はJ-POPが多かった。また、会話そのものへのBGMの影響は強くはなかったが(Fig1)、音楽によって気分や感情が変化することで、意識しない程度に会話も変化している可能性も考えられた。

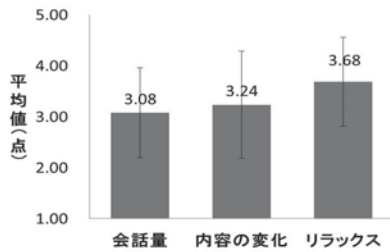


Fig.1: Means and SDs of effects on a conversation and mood by background music.

### 1-10-17

#### 1-10-17 作業可能な環境におけるBGMのシャープネス制御による作業効率向上

Work efficiency improvement by sharpness control of BGM in a workable space  
○星野博之, △上野由美子, △加藤大貴(愛知工業大)

- ◆本研究では、BGMの効果が現れやすい作業課題を検討し、BGMの音響的特徴の一つであるシャープネスを制御することにより作業効率が向上できるのかを調査した。なお、本研究では、RWC研究用音楽データベース(クラシック音楽、ジャズ音楽)を利用した。
- ◆実験1 6曲のBGMを流しながら、3桁+3桁の暗算課題、タイピング課題、就職試験問題SPI3を用いた言語能力課題、を行った結果、暗算課題とタイピング課題で作業効率が評価できることがわかった。
- ◆実験2 クラシック・ジャズから、シャープネス値が異なる3曲ずつ選曲し、さらに、シャープネス値が中間の2曲の高周波成分を減衰または強調させたものを加えた計10曲をBGMに用いて、暗算課題とタイピング課題を行った。Fig.1にシャープネス値とタイピング課題の正タイプ率との関係を示す。結果として、BGMのシャープネス値が大きいほど作業効率(量と精度)が低下傾向となることがわかった。

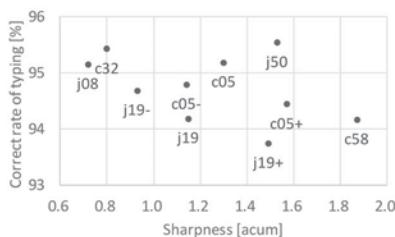


Fig.1 Relation between sharpness of BGM and the typing task results

### 1-10-16

#### 1-10-16 音楽経験と音列の特徴が動的テンポ変化の認知に与える影響

Influence of Music Experience and Characteristics of Tone Sequences on Perception of Gradual Tempo Changes  
○田中陸斗\*, 加藤恒夫\*, 馬田一郎\*\*, 山本誠一\* (\*同志社大, \*\*KDDI 総合研)

- ◆楽器演奏経験及び音列の持つ音楽性がテンポ変化の認知に対する感受性に与える影響について検討した。
- ◆テンポが徐々に遅い方向に変化する音列(deceleration)と、速い方向に変化する音列(acceleration)の音列を聴取し、変化に気付いた時点のBPMと変化する前のBPMとの差を評価指標として取得した。
- ◆音列は、ピアノC4の単一の特徴の音列、ピアノC4の音に4拍子の拍構造を付与した音列、『ツェルニー30番練習曲 第1番』の楽曲の3種類を用いた。
- ◆楽器演奏経験の有無で実験協力者を2グループ(Musician, Non-musician)に分けて分析した。
- ◆Musicianはaccelerationの音列における認知がdecelerationの音列より有意に遅いことが示された。
- ◆Non-musicianはaccelerationの音列とdecelerationの音列との認知に有意差が無かった。
- ◆楽器演奏の練習により、accelerationの音列はdecelerationの音列に比べて認知が遅れる傾向は楽曲以外に単純な音列でも生じることが示されたが、その理由は今後の研究課題である。



Fig.1: Bars from Musical Score of Czerny No.1.

### 1-10-18

#### 1-10-18 ロボットによるBGM認識の検討

A study of background music recognition for robots  
○高橋 徹, △赤塚 俊洋(阪産大)

- ◆ロボット聴覚として聞く対象音を音声から音楽へ拡張したい
- ◆音声対話ロボットのためのBGM認識では收音が困難
  - 近くの音声 → 大きな音量で收音 (クリッピングの問題)
  - 遠くのBGM → 極小音量で收音
- ◆極小音量で收音した楽曲の検索
  - (従来) 混合音「音声+音楽音」で検索するのに適した音響特徴量としてバイナリクロマスペクトル(BCS)を提案してきた
    - ◇ BCS 対 CV (クロマベクトル) BCSが有利
  - (従来) 混合音では有効だったが、極小音量ではどうか示した
    - ◇ BCS 対 CV BCSが有利
  - (今回) 極小音量でCVに音量不変性を導入(正規化)すると?
    - ◇ BCS 対 正規化CV 正規化CVが有利
- ◆496曲で楽曲名と楽曲の再生位置をTop1で推定し精度比較

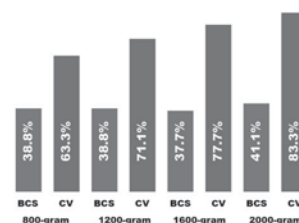


Fig.1: BGM Recognition Accuracy

### 1-10-19

#### 1-10-19 非楽音の打音推定システムの開発

Development of identification system on non-musical percussive sound

☆對馬朗, △伊藤和寿(芝浦工業大学)

- ・本研究では、非楽音である打楽器(ドラムセット)の自動採譜を行うシステムの開発を目的とする。
- ・開発するシステムの最終的な目的は、打音ファイルを入力とし、打楽器を叩いたタイミングを推定し、七種類の打楽器の推定と、三種類の皮物楽器は演奏時の奏法の推定を行う。これらの結果を踏まえて、以下のFig.1に示すようなドラムセットの譜面を出力することである。



Fig.1: Example of drum set musical score.

- ・開発したシステムは以下の三つのプログラムに分けられる。FFT解析プログラムでは、打音ファイルを入力としてFFT解析を行い打音の周波数のピークを検出する。特徴抽出プログラムでは、得られたピークから特徴周波数を得る。楽器推定プログラムでは、特徴周波数と検出ピークを基に楽器・奏法を推定する。
- ・全ての楽器・奏法において10データ用意し、解析・推定を行った。
- ・単音における楽器名・奏法の推定精度は100%であった。一方、推定を行った三種類の和音は推定精度の平均値が86.7%と単音に比べて推定精度が低下した。

### 1-10-21

#### 1-10-21 マルチトラック楽曲におけるGANを用いた大楽節の自動生成

Automatic Eight Bar Phrase Generation Based on GANs for Multitrack Music

☆松浦功一郎, 平林晃, △北原大地(立命館大)

- ◆二種類の敵対的生成ネットワーク(Generative Adversarial Network: GAN)を用いて、大楽節を自動生成する手法を提案する。
- ◆自動作曲システムの実現が困難とされる要因の一つに、楽曲を表現するための時系列データが複雑で長い構造を有することが考えられる。
- ◆Binary-MuseGANのような従来のGANを用いたマルチトラック楽曲の自動生成手法では、4小節からなる小楽節の生成にとどまっている。
- ◆提案ネットワークでは、ランダムノイズから小楽節を生成するネットワークと、その小楽節を入力とし後半楽節を生成するネットワークを組み合わせることで、大楽節の生成を実現する。
- ◆提案手法で大楽節を生成し、アンケートによって得られた主観評価結果をTable1に示す。いずれの評価指標においても、学習データには及ばないものの、従来手法による小楽節を単純に接続した大楽節を上回る結果となった。

Table1 The result of subjective evaluation

	Harmony	Rhythm	Over-all
Dataset	3.83±0.97	4.10±0.77	3.98±0.82
Proposal	3.10±0.94	3.63±0.80	3.25±0.73
Conventional	2.25±0.83	3.08±0.84	2.73±0.71

### 1-10-20

#### 1-10-20 位相を考慮した調波音・打楽器音分離

Phase-aware harmonic/percussive source separation

☆長友健人, 升山義紀, 竹内大起, 矢田部浩平, 及川靖広(早大理工)

- ◆位相を考慮した調波音・打楽器音分離(HPSS)
  - ・調波音の位相変動を除去したIPCスペクトログラムは時間方向に滑らか
  - ・打楽器音のスペクトログラムは時間フレームごとにスパースであることを利用した手法
- ◆従来法は調波音の位相の構造を考慮した凸最適化問題を解くことでHPSSをした
- ◆従来法は打楽器音に対する罰則項の係数λが性能に大きく影響
  - ・λは調波音と打楽器音の割合を調整する
  - ・λの選択によっては性能が大きく悪化する
  - ・λの選択に対して頑健な手法が必要
- ◆本研究では従来法と比べてλの選択に頑健な手法を確立
  - ・SI-SDRが従来手法と比べて向上し、係数λの選択に対して頑健なことを確認

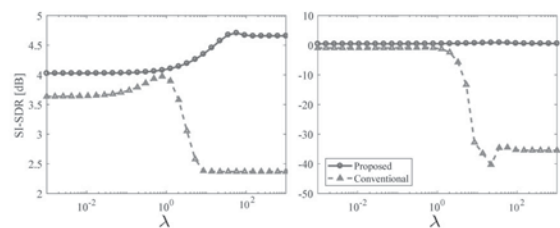


Fig.1 Results of HPSS with each lambda evaluated by SI-SDR (left: harmonic, right: percussive)

### 1-11-1

#### 1-11-1 音響整合材層による圧電高分子トランスデューサについて

P(VDF/TrFE) transducer response using stacking acoustic matching layers

○高橋貞幸(山形大・地域教育文化)

- ◆超音波照射時に圧電振動子と空気との境界で反射が起こり、送受信効率が著しく低下する。圧電高分子振動子と空気での減衰(損失)値は、約68dBと大きいPZTと空気の場合は、更に大きく86dBである。空気中で動作するトランスデューサの効率化を推進するためには、この値をできるだけ小さくすることが必要不可欠になる。
- ◆本研究では、先行研究[1]を基に、音響整合材(ポリマー系)および音響整合層(厚み)を考慮し、効率改善を目的とした、P(VDF/TrFE)トランスデューサの評価を行った。動作中心周波数は、従来通り2MHzとした。本研究で製作した音響整合層付加したP(VDF/TrFE)トランスデューサの概略図をFig.1に示した。この平面型P(VDF/TrFE)トランスデューサは、同じ圧電性能をもつP(VDF/TrFE)トランスデューサより、約10dB効率の上昇が確認された。

・参考文献

[1]高橋貞幸, 音講論(秋), 1129-1130, 2017

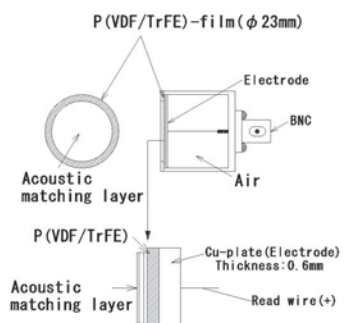


Fig.1: Schematic configuration of the planer type P(VDF/TrFE) transducer with acoustic matching layers (stacking polymers).

### 1-11-2

#### 1-11-2 超音波伝搬に伴い発生するエッジ波の媒質による発生量変化

Change of medium and amount of edge wave generated by ultrasonic wave propagation

◎田中雄介, △吉田光良, △阿部晃, △小倉幸夫(ジャパンブローブ)

- ◆超音波の伝搬に伴い発生するエッジ波は媒質により発生量が異なる。
- ◆エッジ波の発生量は探触子中心軸上において近距離音場限界で大きくなる振幅で評価する。
- ◆エッジ波は空中では発生量が少なく、水やポリスチレンでは発生量が多い。
- ◆媒質自体の減衰である散乱減衰を調べると水や空気は減衰が少なく、ポリスチレンは減衰が多い。
- ◆散乱減衰は近距離音場における最初のピーク位置で評価する。

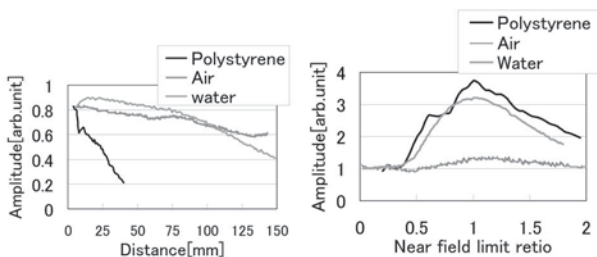


Fig. 1 Relationship of distance and peak A amplitude each medium Fig. 2 Amplitude rising at near field limit by edge wave in each medium (Attenuation corrected)

### 1-11-4

#### 1-11-4 側面が傾斜した生体模擬ファントムの非接触音波伝搬速度計測における精度検討

Study about accuracy of non-contact measurement of speed of sound for a tissue-mimicking phantom with inclined sides

☆大久保 俊朗, 平田 慎之介, 蜂屋弘之(東工大)

- ◆本報告では、空気中から踵側面に超音波を照射し、踵を透過した超音波を検出・評価することで踵内の音波伝搬速度(SOS)を非接触で計測する手法を提案している。
- ◆先行研究で行った側面の傾斜したファントムを対象とした非接触QUS法を用いたSOS計測の精度を検証するため、従来のQUS法を用いた計測との比較を行った。
- ◆QUS法では十分な精度での計測を行うことが出来た。一方、非接触QUS法では、QUS法と比較して計測誤差の影響が大きいということを確認した。今後は、ファントム内を透過する音波の伝搬経路の導出の精度の向上を目指していく。

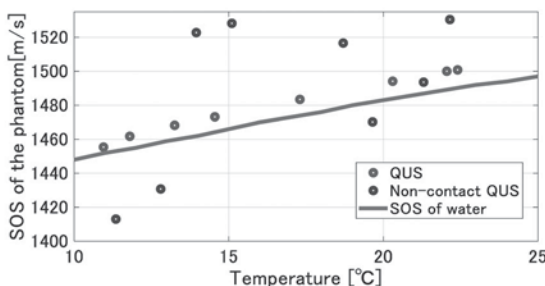


Fig. 1 Estimated SOS of the agar phantom with inclined sides in QUS and in non-contact QUS

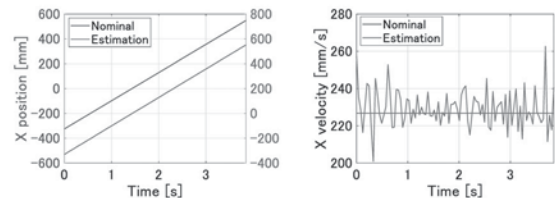
### 1-11-3

#### 1-11-3 M系列信号同時送波による波長より大きな移動物体の位置・速度検出手法の基礎的検討

Measurement of position and velocity for a moving large target using simultaneous transmission of M-sequence ultrasound signals

☆野中 昌吾, 平田 慎之介, 蜂屋 弘之(東工大)

- ◆波長に対し十分大きな移動物体を対象とし、シミュレーションと実験により物体表面における音波の反射点位置が移動する様子を明らかにし、続けて対象物の位置・速度の推定手法について検討した。
- ◆下の条件で、対象物からの反射信号を連続的に取得する実験を行った。
  - スピーカ2個、マイクロホン3個を用いて計測する
  - 2個のスピーカからは異なるM系列信号を同時送波する
  - 対象物としてφ200mmの円柱を速度240mm/sで運動させる
  - 計95フレーム、3.9sの反射信号を取得する
- ◆シミュレーションと実験により、音波が対象物表面で鏡面反射して伝搬していることを確かめ、音波の反射点位置と対象物中心の位置関係が移動に伴って変化することを示した。
- ◆対象物が円柱で近似できる物体の場合、対象物の大きさ及び対象物中心の位置・速度を推定できることが示された。



(a) Target position. (b) Target velocity.

Fig. 1 Estimation results of the target position and velocity.

### 1-11-5

#### 1-11-5 LPMチャープ信号を用いた移動物体の正確な超音波距離計測

Precise ultrasonic distance measurement of moving object by using single Linear-Period-Modulated chirp signal

☆千葉春樹, 黒澤実(東工大)

- ◆チャープ信号の一種であるLPM信号を用いると、移動物体の超音波距離計測が可能である。物体の移動速度に応じて生じる誤差を補償するため、従来手法では速度計測を同時に行う必要があった。
- ◆本報告では、送信信号の解釈を改良することで速度計測を行わずに距離計測誤差を補償可能な手法を提案する。

Fig.1に従来手法、提案手法の計測誤差を示す。提案手法は速度情報を用いずに、従来手法と同等の分解能で計測が可能であり、距離計測値が安定することを確認した。

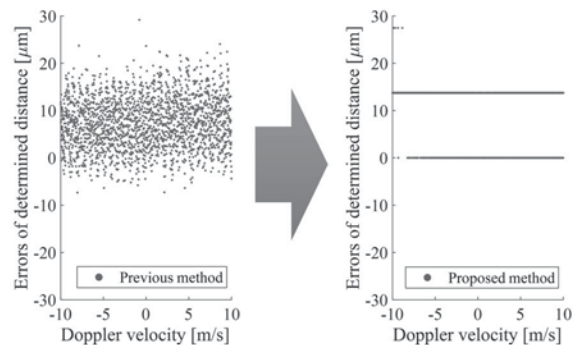


Fig.1: Errors of estimated distance of moving object (Left: Previous method, Right: Proposed method)

### 1-11-6

#### 1-11-6 MUT型音響メタマテリアルによる音響インピーダンス整合に関する基礎検討

Acoustic Impedance Matching by MUT-type Acoustic Metamaterial

○田中宏樹(日立製作所), △町田俊太郎(日立製作所)

近年実用化が進む半導体微細化技術により作製されるMUT (Micromachined Ultrasound Transducer) を利用した音響メタマテリアルによる新たな音響インピーダンス整合を試みた。Fig.1 に提案する基本構造を示す。Acoustic medium から伝わる音波(粗密波)がMembrane structure に到達すると、Membrane が上下に振動し、その際の曲げの力を介して音波がWall structure に伝わる。音響インピーダンスは物質の比音響インピーダンスと伝搬面積の積であるため、 $z_1 \times L = w_1 \times z_0$  とすることで、音響インピーダンス整合が可能となる。提案構造の妥当性をFEMシミュレーションで検証した。水(1.5 MRayl)とシリコン(Si, 20.2 MRayl)の音響インピーダンス整合を評価したところ、従来の1/4入整合層(2層)に対し、MUT型音響メタマテリアル構造は極めて広帯域な透過特性を示した(Fig.2)。

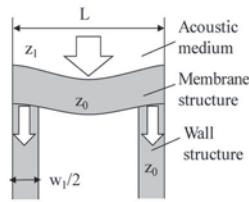


Fig.1: Fundamental structure of MUT-type acoustic metamaterial

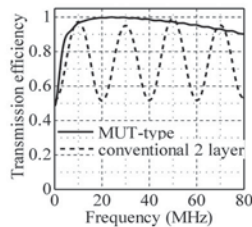


Fig.2: Simulation results (MUT-type and conventional 2 layers)

### 1-11-8

#### 1-11-8 音響キャビテーションが堅牢ハイドロホンに及ぼす影響の検討 - 高速ビデオカメラとレーザーシートを用いた観測 -

Investigation of the influence of acoustic cavitation on tough hydrophone - Observation using a high-speed video camera and a laser sheet -

☆貝瀬不二丸(桐蔭横浜大), 椎葉倫久(日本医療科学大), 森下武志(桐蔭横浜大), 岡田長也(本多電子), 黒澤実(東工大), 竹内真一(桐蔭横浜大)

- ◆当研究室は現在、高強度超音波に50時間曝露しても受波感度が低下しない「前面円柱背面凸型堅牢ハイドロホン」と「前面円錐背面キャップ型堅牢ハイドロホン」の2種類を開発している。
- ◆今回の研究では、より堅牢性を有している前面板の構造を検討するために、高速ビデオカメラとレーザーシートを用いて、音響キャビテーションの様子を観測し、Fig.1のような写真を得ることができた。
- ◆30分程度観測を継続したところ、前面円柱型では継続的にキャビテーションクラウドが付着していた。前面円錐型では、5分程度は付着していたが別のキャビテーションクラウドと合わさり上方へと流れた。その後は観測を終えるまでに1度だけキャビテーションクラウドが飛来したが付着することはなかった。

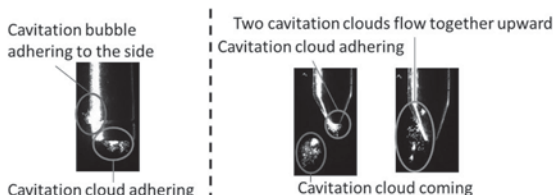


Fig.1: Photographs taken with a high-speed video camera. (Left; Cylindrical titanium front plate tough hydrophone Right; Conical titanium front plate tough hydrophone)

### 1-11-7

#### 1-11-7 堅牢ハイドロホン各部の固有音響インピーダンス、寸法が受波感度の周波数特性におよぼす影響

等価回路を用いた数値計算による検討\*

Effects of specific acoustic impedance and dimensions of each part of tough hydrophone on the frequency characteristics of its receiving sensitivity - Numerical simulation using equivalent circuit-

○竹内真一(桐蔭横浜大・医工), 貝瀬不二丸(桐蔭横浜大・院),

椎葉倫久(日本医療科学大), 岡田長也(本多電子)

- ◆堅牢ハイドロホンの理想的な構造を求めて、MASONの等価回路に基づく一次元音響伝送線路モデルを用いた数値計算により、堅牢型ハイドロホンの前面板、圧電振動子、接着層の固有音響インピーダンスや寸法が受波感度の周波数特性に及ぼす影響について検討した。
- ◆その結果、前面板の寸法の影響で受波感度の周波数特性が一定の周波数ピッチで変動することがわかった。
- ◆また、圧電振動子とチタン製背板の間の接着層の影響により、受波感度の周波数特性の変動幅が大きくなることがわかった。

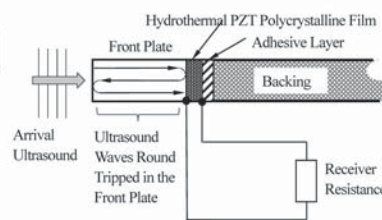


Fig.1: Structure and Concept of Tough Hydrophone.

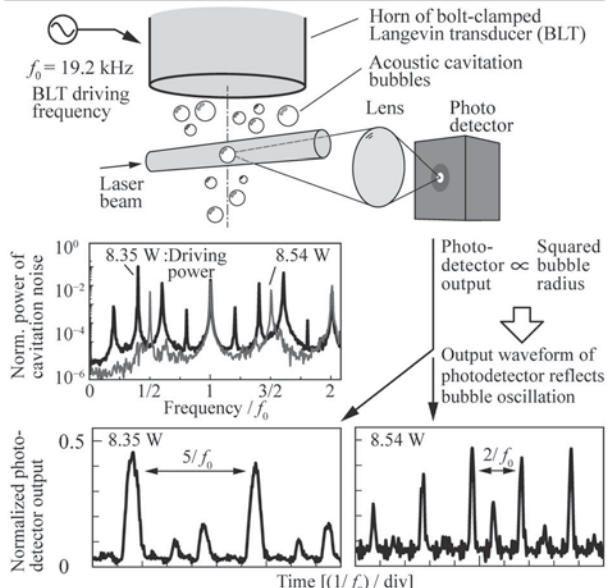
### 1-11-9

#### 1-11-9 超音波ホーン下に生じた音響キャビテーション気泡の分数調波振動

Subharmonic oscillation of acoustic cavitation bubbles under ultrasonic horn

◎黒山 喬允(岐阜高専)

- ・音響キャビテーションノイズと同時にレーザービームの散乱に基づく方法により気泡振動を計測
- ・入射超音波周期の整数倍間隔での気泡分数調波振動を確認
- ・気泡分数調波振動とノイズの分数調波成分に相関
- ・分数調波成分が気泡分数調波振動に起因することを示唆



### 1-11-10

#### 1-11-10 SAW デバイスの高周波化による 小型スイマーの検討

Investigation of a small swimmer by increasing the frequency of SAW devices

☆梶井将, 西尾一輝, 孔徳卿, 黒澤実(東工大)

- ◆弾性表面波 (Surface Acoustic Wave, SAW) デバイスを用いて液中を推進するアクチュエータを SAW スイマーと呼び現在研究を行なっている。先行研究では駆動周波数 9.61MHz のスイマーの駆動に成功している。
- ◆SAW デバイスの特徴として駆動周波数を高くすることによってデバイスの小型化が可能であることから、本研究では駆動周波数を 95.5MHz に上げて小型化したスイマーの検討を行った。
- ◆SAW デバイスを液中に入れた状態でレイリー波を伝搬させると境界面に液体を押し出す力である音響放射力が発生する。音響放射力の反作用によって SAW スイマーは駆動される。
- ◆水面に浮く素材でできたボードに SAW デバイスを取り付けた SAW スイマーを試作し駆動実験を行なった結果、実際に駆動させることに成功した。

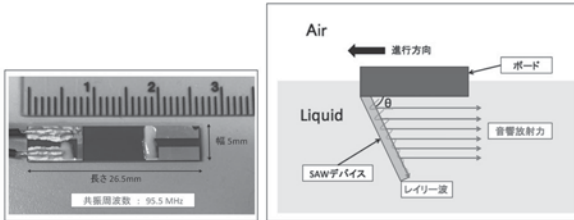


Fig.1: SAW device attached to a swimmer. Fig.2: Driving principle of SAW swimmer.

### 1-11-12

#### 1-11-12 音響インピーダンスを考慮した 皮膚組織の後方散乱係数評価

Backscatter coefficient evaluation of skin tissue considering acoustic impedance distribution

○大村真朗(千葉大・院融合), 吉田憲司(千葉大・CFME), 秋田新介(千葉大・院医), 山口匡(千葉大・CFME)

- ◆後方散乱係数の理論モデルは、音響インピーダンス比の分布と関連するが、実計測信号との対応関係を明らかにする必要がある。
- ◆RF 信号 (15 MHz) と音響インピーダンス分布 (80 MHz) から皮膚組織の後方散乱係数 ( $BSC$ ,  $BSC_z$ ) をそれぞれ算出し、比較した。
- ◆RF 信号ではリフレクター法により  $BSC$  を、音響インピーダンス分布では 2 次元フーリエ変換に基づく理論式から  $BSC_z$  を算出した。
- ◆両者の空間分解能の違いを考慮し、 $BSC_z$  にガウシアンフィルタ ( $\sigma = 3$ ) を適用し、カーネルサイズの違いによる変動を確認した。
- ◆Figure 1 に  $BSC$  と  $BSC_z$  の周波数依存性を示す。両者の相対的な差はあるが、 $LE(-)$  より  $LE(+)$  の方が低い関係は同じである。また、フィルタのカーネルサイズを大きくすると、 $BSC_z$  の傾きが正に近づいた。
- ◆傾きに変動がみられた点について、フィルタリングの効果により、音響インピーダンス分布の局所のばらつきが平滑化され、高周波対低周波成分のパワーが小さくなるためだと考えられる。

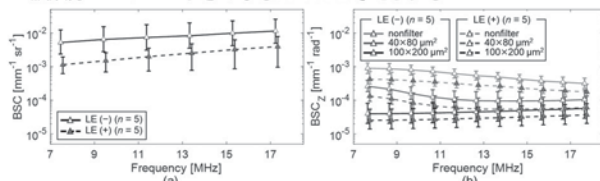


Fig.1: Frequency dependence of measurement (a) and predicted (b) BSCs.

### 1-11-11

#### 1-11-11 4リングアニュアラレイトランスデューサ による音響放射圧の発生と変位の評価

Evaluation of Acoustic Radiation Force generation and displacement by a 4-ring annular array transducer.

☆窪田溪之, 田中直彦(芝浦工大・院)

- ◆近年、エラストックイメージングと呼ばれる超音波検査法が普及してきた。本研究では、超音波顕微鏡が対象としているようなマイクロなスケールにおいて、音響放射圧を用いてエラストックイメージングを行うための基礎的な検討を行う。
- ◆実験は、Fig.1 のような装置を用いて行った。音響放射圧はアニュアラレイトランスデューサによって発生させる。反射体の変位は、50MHz の超音波顕微鏡用のトランスデューサによって計測する。
- ◆アニュアラレイトランスデューサを用いて音響放射圧を加えることにより、反射体に変位することを実験的に確認した。さらに、アニュアラレイトランスデューサによる音響放射圧の分布が概ね予想通りになっていることも確認できた。
- ◆生体組織に近い豚ロース肉を用いた実験では、硬さの違う部分で変位の違いが観測できた。

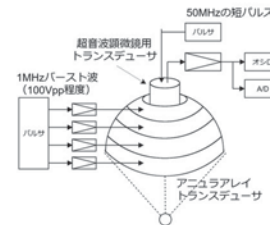


Fig.1 Configuration of experimental equipment

### 1-11-13

#### 1-11-13 10-MHz 帯超音波の音響放射力で 駆動されるマイクロバブルの動態解析

Analysis of translation of microbubble driven by acoustic radiation force of ultrasound with frequency of more than 10 MHz

○吉田憲司(千葉大・CFME), 齋藤勝也, 大村真朗(千葉大・院融合), 山口匡(千葉大・CFME)

- ◆リンパ管のイメージング方法として、音響放射力により駆動されるマイクロバブルの移動量を定量する動的造影超音波法を提案している
- ◆ファントム中に形成した模擬リンパ管に造影剤を充填し、中心周波数 14.4 MHz の超音波を繰り返し周波数 2 kHz で照射した際のマイクロバブルの動態を受信エコー信号の解析を基に実施した。
- ◆Figure 1 は模擬リンパ管の B mode 画像、管中央部分における M mode 画像および造影剤移動速度の時間変化である
- ◆超音波照射直後では管全体 ( $z=1.93\sim 1.99$  mm) で造影剤の移動が検出されているが、時間経過とともに特に管壁部分での移動が検出されなくなる様子が確認された
- ◆管内の造影剤の動態は過渡的であり、この特徴を考慮した超音波送受信パルスシーケンスを考案していく必要が示唆された

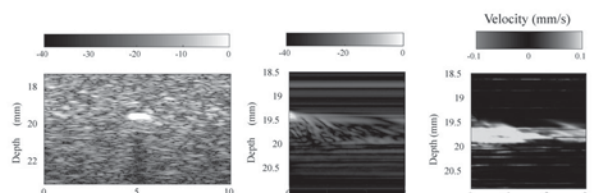


Fig.1: Transient behavior of ultrasound contrast agents in chamber with 0.5 mm diameter. (a) preliminary obtained B mode image, (b) M mode image and (c) time variation in translational velocity.

### 1-11-14

#### 1-11-14 複素数空間での振幅確率密度関数モーメントを用いた肝線維化パラメータの推定精度評価

Evaluation of estimation accuracy for parameters of liver fibrosis by using moments of the probability density function of echo envelope in complex space

☆張闓, 平田慎之介, 蜂屋弘之(東工大)

超音波画像を用いた肝線維化定量評価を目的とし、線維化肝組織の超音波画像の振幅確率分布モデルであるマルチレイリーモデルに基づく肝線維化パラメータ推定手法を提案している。先行研究では、複素数空間のすべてのモーメントを探索しておらず、さらに、推定精度の評価値としてモーメントデータの95% (±2σ) が含まれる領域の境界線をモデルパラメータ空間に射影した結果を用いていた。しかし、射影時の誤差により、過剰に評価する傾向があるため、新たな評価方法も検討した。本報告では、標準偏差と±σ境界による推定精度評価方法を用いた結果を示す。その結果、複素数空間で1.55次(実部)-0.65次(虚部)付近のモーメントを用いると肝線維化パラメータを高精度に推定できる見通しを得た。Fig. 1は、外の境界線はデータの95%が含まれる領域の境界、中の境界線はデータの65%が含まれる境界、点は1000枚の超音波画像の推定結果を示している。

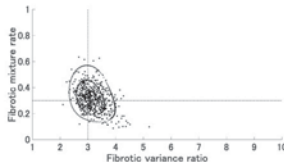


Fig. 1 Distribution of fibrotic parameters space using projection of contour line using 1.55th and 0.65th moments in real and imaginary spaces. The setting fibrotic parameters, mixture rate and variance ratio are 0.3 and 3.

### 1-11-16

#### 1-11-16 生体内を伝搬するずり弾性波による疑似境界波の発生について

The quasi-boundary wave generated by a shear wave in biological tissues

○佐藤雅弘(富山大・工)

- ◆最近、肝臓内を伝搬するずり弾性波(横波)の位相速度を計測して肝硬変や肝がんなどの早期発見を目的とした研究が行われている。
- ◆先に筆者らは、上記に資する目的で、数値シミュレーションにより、横波の伝搬や病変部を通過するときの波動のふるまいを解析した。
- ◆解析の結果、励振の際に発生する縦波は周囲が自由境界であっても、その伝搬は観測されなかった。また、病変部と正常部の境界に境界波と思われる横波・縦波成分が観察されたが検証には至らなかった。
- ◆弾性波動の反射・屈折特性において、遅い媒質から速い媒質に平面波が入力する時、臨界角付近で一種の境界波が発生することが知られている。そこで、上記シミュレーションの媒質条件で反射・屈折特性を解析したところ、同様に境界波が発生することが分かった。
- ◆それを検証するため数値シミュレーションを行った結果、境界波の存在が確認された。

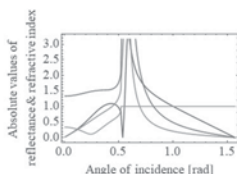


Fig.1: Absolute value of reflectance and refractive index vs angle of incident plane shear wave. Pi, Si, Pr and Sr show reflected P, reflected SV, refracted P and refracted SV waves, respectively.

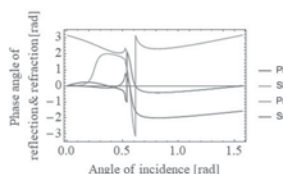


Fig.2: Phase angle of reflective and refractive waves vs angle of incident plane shear wave.

### 1-11-15

#### 1-11-15 心腔内血流計測のための特異値分解クラッタフィルタの空間分割化に関する検討

Investigation of depth-dependent singular value decomposition clutter filter for cardiac blood imaging

☆茂澄倫也, 長岡 亮, 長谷川英之(富山大)

◆近年、特異値分解(SVD: singular value decomposition)に基づくクラッタフィルタが注目されている。しかしながら、心臓は深さ方向に時空間特性が異なるため、深さごとにクラッタフィルタを用いることが望ましい。本報告では、心腔内血流画像の取得のために特異値分解フィルタを深さごとに適用することを検討した。

◆ Fig. 1(a)および Fig. 1(b)に従来法および提案法によって取得された左心室内腔の血流のエコー画像を示す。Fig. 1より、提案法により浅部領域における血球エコー信号の見えが改善されていることが読み取れる。

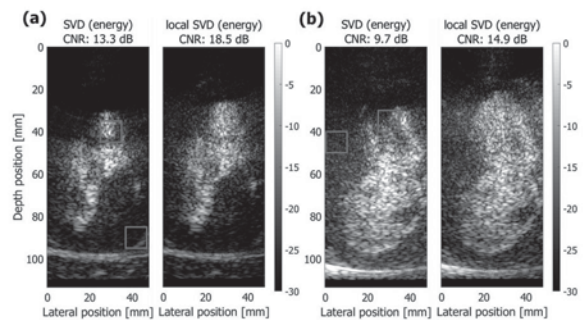


Fig. 1. (a), (b)収縮期および拡張期における血球エコー画像

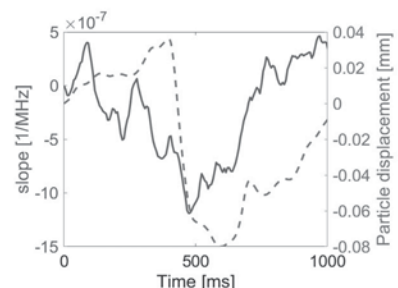
左: 従来法, 右: 提案法

### 1-11-17

#### 1-11 高時間分解能超音波イメージングを用いた血球の粒子サイズと圧力との関係に関する基礎検討

◎長岡 亮\*, 大村 眞朗\*, 山口 匡\*, 長谷川英之\*(富山大), \*千葉大)

平行波や拡散波を用いた高い時間分解能を有する超音波計測法が確立し、生体内の高速で移動する血液のダイナミクスの可視化が可能となってきている。赤血球は加わる圧力や流速によって、その径や凝集度に変化すると考えられ、血流エコーから赤血球の粒子サイズが推定出来れば、その原因である圧力や流速を間接的に評価することが可能となる。本報告では、血流エコーの周波数スペクトルから、粒子サイズと逆数の関係にあるその傾きを算出し、血液の流速と血管壁の粒子速度とをそれぞれ比較した。特に、圧力と線形的な関係がある血管壁の粒子変位に関しては、血流速よりも波形の立ち上がりのタイミングが一致しており、より強い関係性があるのではないかと考えられる(下図)。



Relationship between the estimated slope and particle velocity of the artery wall.



### 1-11-18

#### 1-11-18 地中埋設物の三次元映像化 -表面波モデルについて-

Three dimensional underground imaging  
-A study on the model of surface wave-  
☆泉田健吾(千葉工大), 陶良(千葉工大)

地中埋設物の非破壊探査技術として、電磁誘導型音源と振幅相関合成処理を併用した映像化法の研究が進められてきた。しかし受信信号に埋設物の反射波と共に表面波が強く含まれ、この表面波成分を除去することは映像化精度の向上を期待できる。

本稿では表面波の推定手法として実際の音響特性を想定した周波数領域のLPFモデルを考案し、従来の時系列モデルを用いた推定結果との比較検討を行った。表面波の推定では、Levenberg-Marquardt法を用いて受信波形に対しモデルのフィッティングを行い、その残差を表面波推定効果の評価とした。

96個の実測信号を用意し、従来のモデルと今回考案したモデルの残差率を比較した結果をFig.1に示す。ほぼ全ての信号において周波数領域のLPFモデルがより低い残差率となった。

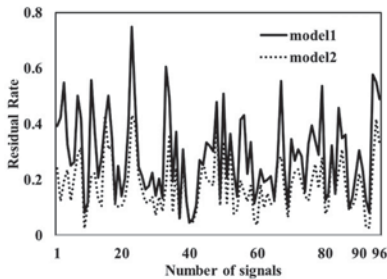


Fig.1 Comparison of Residual Rate

### 1-11-20

#### 1-11-20 強力空中超音波波源走査法による高調波ガイド波伝搬を利用した非破壊検査の検討

Investigation of non-destructive inspection using harmonic guided wave propagation by scanning high-intensity aerial ultrasonic source technique  
☆山田健太, 大隅歩, 伊藤洋一(日大・理工)

- ◆点集束強力空中超音波(40kHz)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、音波の非線形性により発生する高調波を利用して、複数周波数の音波伝搬を可視化し、その伝搬情報から欠陥の検出を試みた。
- ◆その結果、周波数によって減肉欠陥周辺の音波伝搬に変化があることが分かった。また、音波受信位置を変更しても高周波を利用することで、減肉欠陥の検出に成功した。

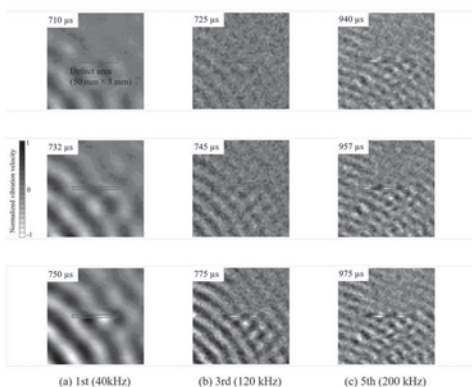


Fig.1 Distribution of vibration velocity

### 1-11-19

#### 1-11-19 音響特性による室内温度の推定 -音響特性の計測による室内状態の検知(II)-

-Estimation of room temperature through acoustic characteristics-  
Detection method of room conditions by measuring acoustic characteristics (II)  
☆林田耀生, △水野洋輔, 中村健太郎(東工大)

- ◆部屋の音響特性の変化から室温とその分布を推定することを目的として、1/8縮小モデルを用いた基礎実験を行った。
- ◆縮小モデル内に置いたスピーカからチャープ音を出力し、マイクロホンで收音し、1Hz~20kHzの範囲で周波数応答を測定する。
- ◆室温の推定は、温度変化により周波数応答形状が周波数軸方向に音速の比で伸縮することを利用する。
- ◆室温を常温付近の範囲で変化させ、室温推定を行った。精度の検証のためにIC温度センサを設置し比較したところ、0.5°C程度以内の誤差で推定できた(Fig.1)。
- ◆温度分布がある場合についても実験を行い、温度分布の大小の推定が可能であることを示した。

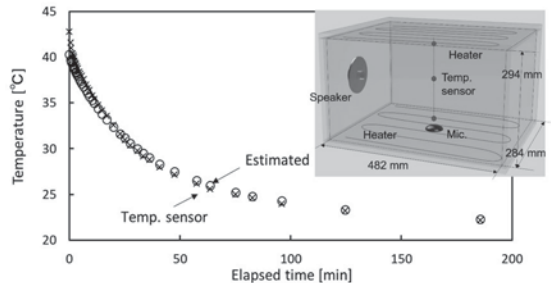


Fig. 1 Temporal variations of actual and estimated temperatures (within 0.5 degrees of accuracy).

### 1-11-21

#### 1-11-21 最小二乗法を用いた可視化音場からのインテンシティ推定 Intensity estimation using the least squares method for the visualized sound field

☆今枝文彦, 矢田部浩平, 及川靖広(早大理工)

- ◆2次元可視化音場からのインテンシティ推定法を提案
- ◆従来よりも高次の多項式モデルで近似するかたちに拡張
- ◆周波数に関する数値シミュレーションを行い従来法と比較、従来よりもノイズに強く、性能が向上することを確認

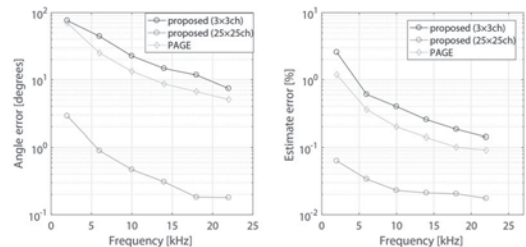


Fig.1 Estimation error for each method when using pressure gradient

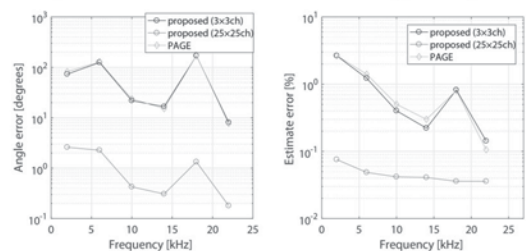


Fig.2 Estimation error for each method when using phase gradient

### 1-11-22

#### 1-11-22 ストロボ効果を利用した PIV による 粒子速度分布の測定

PIV measurement of particle velocity distribution using stroboscopic effect

©佐脇真平, 土屋裕造(戸田建設)

- ◆PIV (粒子画像流速測定法) による非接触での音場計測法の提案。
- ◆測定対象を単一周波数が卓越した定常音場に限定し, 照射する光シートを点滅させることで, ストロボ効果により空気の見かけの周波数を小さくする。目視での観察に加え, 高い感度および大きな画素数を有するカメラを用いた PIV で粒子速度の測定を試みる。
- ◆音響管実験により粒子速度分布が測定できることを確認。

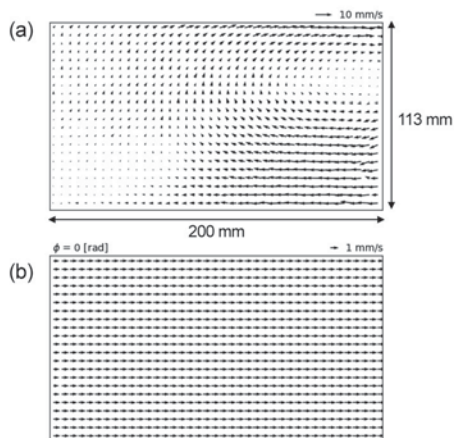


Fig. 1 (a)Flow velocity distribution and (b)particle velocity distribution inside the sound tube.

### 1-11-24

#### アニュラアレイによる生体組織の 後方散乱特性評価の高精度化

Improvement of backscattering property evaluation of soft tissue by annular array

○溝口岳(千葉大院・融合), 田村和輝(浜松医大), Jonathan Mamou, Jeffrey A. Ketterling(Riverside Research), 吉田憲司, 山口匡(千葉大・CFME)

- ◆中心周波数 20 MHz のアニュラアレイで正常肝モデルの摘出ラット肝臓を計測し, 散乱体径の異なる(10, 20, 30 μm)ファントムのエコー信号をそれぞれリファレンスに用いて後方散乱係数を評価した。
- ◆参照媒質の条件により, 推定される後方散乱係数が異なることが確認された。特に散乱体径 10 μm のファントムで標準偏差が最小かつ周波数依存性の低い結果となった。(Fig. 1)
- ◆正常肝モデルのラット肝臓では 10 μm 程度の組織が主な散乱体であると考えられる。今後は疾患モデルラットの肝臓での評価を行う。

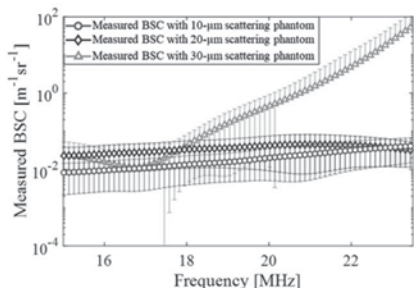


Fig. 1: The BSC curves of the rat liver

### 1-11-23

#### 1-11-23 マルチスケール音速解析による肝臓組織 の構造特性評価

Structural characterization of liver by multi-scale speed of sound analysis

○小川拓朗(融合理工, 千葉大), 吉田憲司(CFME, 千葉大), 松本大輝(工, 千葉大), 大西俊, 羽石秀明, 山口匡(CFME, 千葉大)

- ◆正常モデルラット肝臓の薄切試料を対象とし, 分解能の異なる振動子(中心周波数 60, 80, 250 MHz) を広域 (5.0×3.6 mm<sup>2</sup>) で走査し, ミクロサイズ (1 pixel = 2×2 μm<sup>2</sup>) の情報量を保持しながら試料全体の組織構造を踏まえた音響特性解析を行った。
- ◆計測後の試料を HE 染色後, それぞれの解析結果に対して病理像との位置合わせを行うことで同一箇所における音響特性の比較を行った。
- ◆肝細胞の密集部分 (200×200 μm) において, 60, 80, 250 MHz での観察を比較すると, それぞれの音速の平均と標準偏差は 1672.6±28.7, 1717.6±47.7, 1726.6±77.0 [m/s]となり, 周波数が高くなるほど平均および標準偏差が高値であった。
- ◆分解能が低くなると, 複数の組織の平均特性として評価されることが想定される。

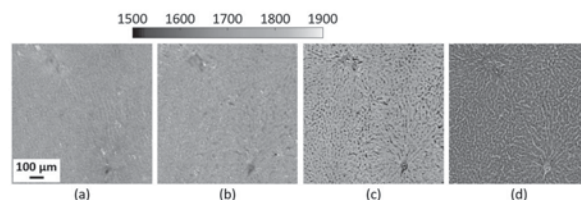


Fig. 1: 2D speed of sound map (1.0×1.0 mm<sup>2</sup>) of control liver evaluated with center frequency of 60 MHz (a), 80 MHz (b), 250 MHz (c), and histological image (d)

### 1-P-1

#### 1-P-1 エンジン音振分析の応用による音響特微量 の可視化手法

Visualization method of acoustic features by application of engine NV analysis

○今井将之(サイエンスカフェ DEMEMS), △宮代大輔(東理大)

- ◆アニメの声優は役によって声色や口調を使い分けているが, それでもアニメ好きはキャラクターの声から声優を当てることができる。これは各声優が固有の声質を持つためである。その声質を調べるためスペクトログラムや包絡スペクトルで分析すると, アニメのセリフは時間変化するため, 定量的に分析するのが困難であった。そこで今回, エンジン音振解析を応用してアニメのセリフから音響特微量を定量的に可視化する手法を考案した。
- ◆エンジンの音振解析の1つにエンジンの回転数毎にスペクトル解析を行い3次元マップで表現するrpm スペクトラム・マップがある。
- ◆この回転数を音声の基本周波数(F0)に置き換え, 時間軸のスペクトログラムから F0 軸に変換することで, 声優の声質を定量的に比較することができた Fig. 1。本会ではその詳細を発表する。

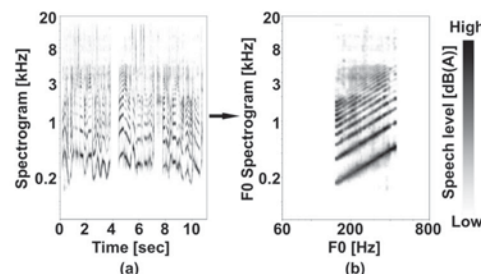


Fig. 1 Transform from time axis to F0 axis

## 1-P-2

### 1-P-2 既知の工場環境音を用いた深層学習に基づく工作機械雑音除去の検討

Deep Learning-based Noise Reduction using Known Machine Sound in a Factory

☆レオ チーシャン, 西崎博光(山梨大・院)

◆研究目的:

深層学習(オートエンコーダ)を用いて、音声収録する場所の雑音環境に適応した雑音除去方法を検討

「厳しい実雑音環境下でも従来手法がどれだけ頑健なのか?」を検証

◆対象音声:

機械加工工場内で実際の技術者が作業しながら収録した音声

➢ 様々な工作機械の音、雑音の大きさも時間毎に変化

◆雑音除去モデル:

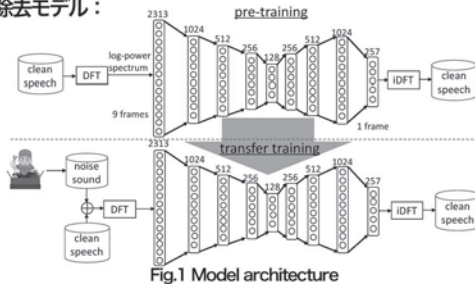


Fig.1 Model architecture

◆実験結果

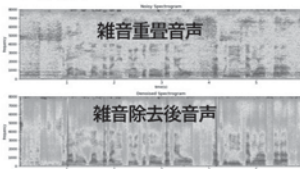


Fig. 2 Log-power spectrums

Table1 MOS and ASR results.

	MOS	W. Corr. [%]
before	2.7	59.7
after	3.2	62.0

厳しい雑音環境下においても、主観評価(MOS)、客観評価(W. Corr.)の双方で改善

## 1-P-4

### 1-P-4 音響・言語モデルの同時適応による感情音声認識の精度改善

Improvement of emotional speech recognition by simultaneous adaptation of acoustic and language models

☆佐伯和哉, 加藤正治, 小坂哲夫(山形大), 能勢隆(東北大)

◆感情音声認識では、感情の種類や感情強度による音響特徴の変動に加え、言語的特徴が平静音声と異なるため認識を難しくしている。本研究では感情音声認識の精度向上を目指し、音響モデルおよび言語モデルの同時適応を行う。

◆認識対象として感情音声コーパス JTES (Japanese Twitter-based Emotional Speech)を利用する。このコーパスはTwitterの呟きをベースに、4感情(怒り・悲しみ・喜び・平静)に該当する文が、音韻・韻律のバランスを考慮して選択されている。

◆認識実験では、音響モデル適応、言語モデル適応単独の場合と両者を併用した場合で比較を行った。結果を表1に示す。

◆実験より、同時適応は効果が高いことが分かった。また認識結果の分析から言語モデル適応が効果的に働くのは文末に多く見られることが分かった。

Table.1: Word error rate for each model adaptation (%)

	Baseline	AM adapt	LM adapt	AM+LM adapt
Ang	36.53	23.61	29.72	22.92
Joy	39.08	30.62	30.15	21.39
Neu	21.81	18.19	19.52	17.71
Sad	38.48	24.24	32.12	21.67
Ave	33.97	24.17	27.88	20.92

## 1-P-3

### 1-P-3 耐雑音音声認識の雑音・無雑音ウェーブレットスペクトルの圧縮法の基礎的検討

A study of noisy speech recognition using the compression method on wavelet spectrum for clean and noisy speech data

○緑川洋一, 秋田昌憲(大分大・理工)

雑音環境下における音声認識の認識率の改善は、音声入力装置の一般への普及のためには不可欠な技術のひとつである。しかし信頼性のある決定的な解決方法と言えるような技術は現段階ではなく、重要な解決課題の一つであると考えられる。我々はこのような点に着目し、雑音環境下における音声認識について以前より研究を行ってきた。一般的に信号処理で用いられる解析手法のフーリエ変換による周波数領域における処理とは異なるウェーブレット変換を用いることにより雑音環境下音声認識の認識率改善ができないか基礎的な検討を行っている。

そのような研究の一端として、我々は雑音環境下の音声認識の認識率改善のため、認識時の標準データとする無雑音の音声データにはウェーブレット変換を用いてデータの一部を強調することにより擬似的に雑音をのせるような処理をして、入力した雑音の重畳した音声データにはウェーブレット変換を用いてデータの一部を圧縮することにより雑音の低減をし、両者を併用することにより特徴を近づけ、認識率の向上を狙う方法における圧縮の基礎的検討を今回は行ったので報告する。

## 1-P-5

### 1-P-5 Automatic Generation of Text with Errors using Neural Machine Translation for Grammatical Error Detection

○Jiang Fu, Yuya Chiba, Takashi Nose, Akinori Ito (Tohoku Univ.)

◆The speech-based CALL systems based on automatic speech recognition (ASR) technology have been developed.

◆Grammatical error detection (GED) from speech is a task to detect grammatical and lexical mistakes from spoken sentences.

◆We introduced the generation of text with grammatical errors by using NMT in the GED system (Fig.1).

◆By using the three different language models, the contents containing more related grammatical errors will help to improve the speech recognition accuracy in the GED system (Fig. 2).

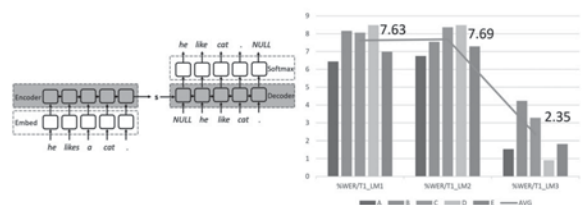


Fig.1: A word level NMT for generating text with grammatical errors

Fig.2: Speech recognition results by using three language models

### 1-P-6

#### 1-P-6 日本語の Posteriorgram を用いた ゼロ資源言語の音声検索の検討

Spoken Term Detection Using Japanese Language Posteriorgram  
for Zero-Resource Language

☆水落 智, 千葉 祐弥, 能勢 隆, 伊藤 彰則(東北大学)

- ◆本研究では、ゼロ資源言語のデータベースに対する単語検索として、日本語の音素識別器から得た Posteriorgram を用いる手法を検討した。検索対象言語はカクチケル語とする。
- ◆音素認識器の入力には、当該フレームの MFCC に加えて、前後数フレームの MFCC を結合したものを用いる。
- ◆異なる言語の音素認識器から得られた Posteriorgram を用いると、MFCC と比較して検出性能を改善することができたが、その改善幅は同じ言語による Posteriorgram を利用する場合に比べて小さいものであった。
- ◆今後は言語の差異による影響を低減するため、複数言語を用いた特徴量抽出に関して検討する予定である。

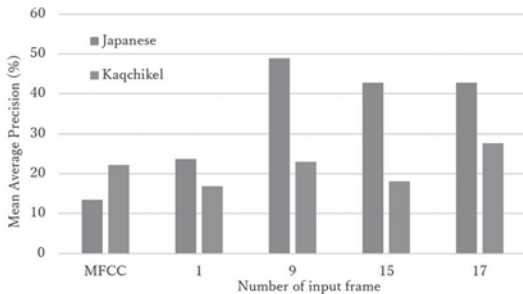


Fig. 1: MAP of STD using Japanese language Posteriorgram

### 1-P-8

#### 1-P-8 複数の音声認識結果表現を用いた 系列変換モデルによる音声翻訳システムの検討

A study on speech translation systems using various representation of ASR results

☆民谷慎一郎, 秋葉友良, 塚田元(豊橋技科大)

- ◆本研究では、End-to-End 音声認識モデルと End-to-End 機械翻訳モデルを用いた音声翻訳システムを対象とする。中間表現である音声認識結果に書記素、音素の2通りを用いる複数表現と、異なるサブワード単位を用いる複数サブワード単位でアンサンブル音声翻訳をして、認識結果を多様化することによる音声認識誤りの影響の低減を検討する。
- ◆英独の音声翻訳実験の結果、複数表現、複数サブワード単位のアンサンブル音声翻訳モデルが書記素の音声認識結果を用いて同一サブワード単位で翻訳を行うベースラインモデルの性能を改善することを確認した。

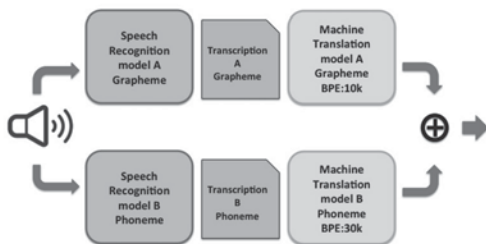


Fig. 1 ensemble speech translation model

model	BLEU
baseline	13.99
multi-subword	14.60
multi-representation	14.18

Table. 1 speech translation result

### 1-P-7

#### 1-P-7 VAD 入出力間の類似性比較による 基本周波数推定

F0 estimated from a similarity between input and output features of DNN-based VAD

☆信楽希, 山森圭(金沢大),

△Dwijayanti Suci(Sriwijaya 大), 三好正人(金沢大)

- ◆音声の基本周波数(F0)は、音の高さを表す特徴量であり、様々な音声信号処理で重要とされる。私たちは、推定性能の高い音声区間検出(VAD)を応用することで、雑音環境下でのF0推定精度向上を目指している。
- ◆応用を検討中のVADは、5層の制限付きボルツマンマシン(RBM)からなるディープニューラルネットワーク(DNN)を用いている。DNNへの入力特徴量は、Log power spectra(LPS)と入力音声の中の音声区間を強調した独自特徴量 Speech period candidates(SPCs)である。
- ◆SPCsとVAD結果の類似性を日本語音声のモーラ長程度の時間間隔で比較し、類似性の高い上位5サブバンドを選択することでF0の推定を行った。Fig.1は、ホワイトノイズ0dBにおける推定結果である。(選択されたサブバンド: 淡青, F0を含むサブバンドのうち上位5に含まれるもの: 黄・含まれないもの: 赤)
- ◆Fig.2は、上位5サブバンドF0の数を、ノイズの種類・SN比毎に示す。類似性をもとにサブバンドを選択することで、平均で84%のF0を推定することができる。
- ◆類似性比較によって選択されたサブバンドをF0推定であるSWIPEへ適用した。変更を加えたSWIPEについて、F0推定精度を議論する。

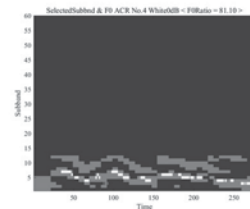


Fig. 1: Estimated F0 of a speech signal smeared with White noise (0 dB)

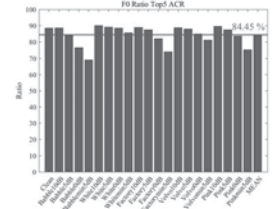


Fig. 2: Number of F0s included in top 5 subbands

### 1-P-9

#### 1-P-9 双方向 LSTM に基づくオンライン認識における課題

A Study on Bidirectional-LSTM-Based End-to-End ASR for Low-Latency Decoding

○小林 彰夫(筑波技大)

- ◇低遅延 end-to-End 音声認識
  - 聴覚障害者の発話を対象として、低遅延 end-to-end 音声認識の課題を調査
  - オンライン向け低遅延の end-to-end 音声認識では、一定の連続フレームを対象に LSTM を動作させる [Zang et al., 2016]
    - 発話途中からの LSTM の遷移を行うため、コンテキスト推定が不十分
    - 聴覚障害者は発話が不明確になりやすい⇒認識性能の劣化が予想
- ◇実験
  - 学習データ: JNAS, 評価データ: S-JNAS, 聴覚障害者の発話(バランス文 B/C セット)
  - ネットワーク: 4層 LSTM(双方向/単方向), CTC またはアライメントに基づく学習, 文字出力
- ◇実験結果
  - 聴覚障害者(オフライン): 双方向 LSTM(CTC(Bi)) を使っても性能が低く、順方向 LSTM のみでは性能が著しく低下する
  - 聴覚障害者(低遅延): 性能が低く、逆方向 LSTM によるコンテキストの予測が困難。性能改善に必要なチャンクサイズの増加の問題を解決する必要がある

表 1: 実験結果(オフライン)

	S-JNAS	聴覚障害者	
		男性	女性
CTC(Bi)	0.4	25.8	13.0
CTC(Uni)	0.6	25.7	15.2
CE(Bi)	3.1	72.7	27.7
CE(Uni)	6.6	90.1	61.5

表 2: 実験結果(低遅延)

チャンク/シフト	S-JNAS	聴覚障害者	
		男性	女性
6/3	9.8	92.1	64.7
10/5	6.1	82.5	50.5
15/5	5.5	83.3	42.6
20/20	4.5	76.4	40.2

### 1-P-10

#### 1-P-10 Speech analysis using the second derivative of phase spectrum

☆Nguyen Binh Thien (Ritsumeikan Univ.), Yukoh Wakabayashi (Tokyo Metropolitan Univ./Ritsumeikan Univ.), Takahiro Fukumori and Takano Nishiura (Ritsumeikan Univ.)

Instantaneous frequency (IF) and group delay (GD) are the two important representations of phase spectrum. They play the major role in phase-aware speech signal analysis. IF tracks the fluctuation of the phase along the time axis, while GD shows the change of the phase with respect to frequency. This paper proposes a new representation of phase spectrum based on the ideas of both the GD and the IF, i.e., the second-order mixed partial derivative of the phase (SDP). The result shows that the new representation has advantage to reveal the structure of speech signal.

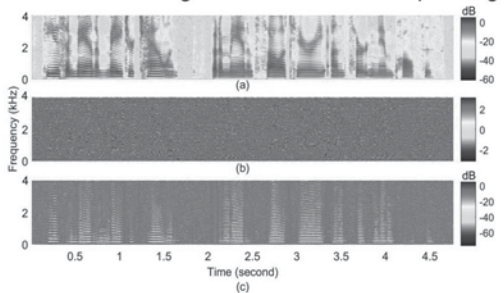


Fig. 1 Representations of a clean speech signal in time-frequency domain, using Hann window with 32 ms duration, 4 ms frame shift, and 4096-point FFT with zero padding. (a) Amplitude spectrogram. (b) Phase spectrogram. (c) SDP in logarithmic scale.

### 1-P-12

#### 1-P-12 講演音声認識の修正語のオンライン教示による半自動的な修正手法と語彙適応の併用の効果

Semi-automatic correction and vocabulary adaptation by online feedback of correct words in speech recognition for lectures

△寺田侑司, ☆塚本皓斗(静岡大院・総合科学技術研), 甲斐充彦(静岡大)

- ◆本稿では自動字幕を想定した音声認識の修正支援システムにおいて、修正語のフィードバックによる字幕の半自動的な修正と言語モデル適応の併用の効果を検証している。
- ◆修正語を追加する方法とベース言語モデルの学習データから修正語を含むテキストを自動生成する方法を組み合わせ、単純に半自動修正字幕で適応した場合と比べ単語誤りを11.5%削減した。
- ◆適応した言語モデルとLSTM言語モデルを組み合わせることにより、未知語の誤りを大幅に減少した。

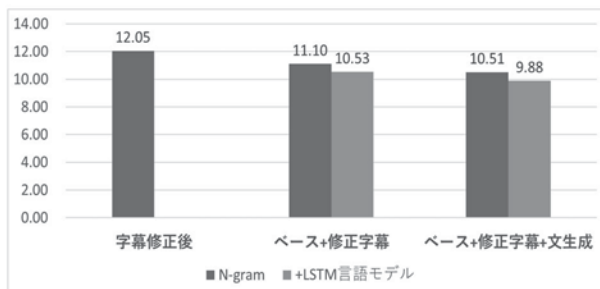


Fig.1: CSJ core11講演中の全発話の単語誤り率[%]

### 1-P-11

#### 1-P-11 講義音声認識のための回り込み音声の影響分析とDNN音声分離モデルによる改善の一検討

Influence analysis of wraparound speech for speech recognition in lecture room and an investigation on improved ASR using DNN-based speech separation frontend

△脇屋義也(静岡大院・総合科学技術研),

☆福井明日香(静岡大院・総合科学技術研), 甲斐充彦(静岡大)

- ◆講義室の環境において音声認識の精度が低下する原因についての分析はまだ不十分である。
- ◆講義環境を模擬した音声収録を行い、おもに拡声機を通して混入する音声が発話音声認識に与える影響についての分析を行った。
- ◆収録マイクの違いによる影響の違いや、対象話者(講師)以外の音声の混入だけではなく、本人の音声の混入が発話音声認識に悪影響を与えることを明らかにした。
- ◆DNN音声分離モデル適用により改善効果が得られることを示した。

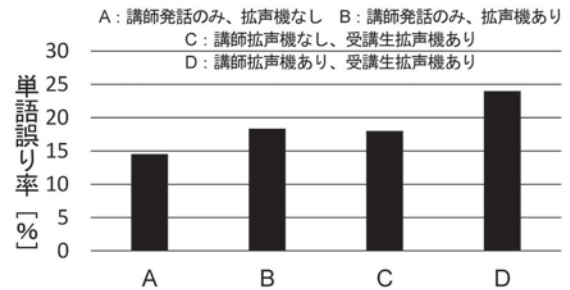


Fig.1: 混入音声の有無と拡声機の影響の比較

### 1-P-13

#### 1-P-13 構音障害者を対象とした日本語大語彙連続音声認識の検討

An investigation of Japanese large-vocabulary continuous speech recognition for dysarthric speakers

◎高島遼一, 滝口哲也, 有木康雄(神戸大)

- ◆アテトーゼ型脳性麻痺による構音障害者の音声認識について検討する。本構音障害者と健常者の発話スタイルは大きく異なるため、健常者音声で学習した従来の不特定健常者音声認識モデルでは認識が困難である。
- ◆本稿では、不特定健常者モデルからのモデル適応により、特定構音障害者モデルの構築を試みる。ただし前述の通り、発話スタイルの差が大きいため、不特定健常者モデルから直接、特定構音障害者音声へ十分な精度で適応することは困難と考えられる。
- ◆本研究では不特定健常者モデルから、不特定構音障害者モデルへの適応を介して特定構音障害者モデルへ適応するという2段階のモデル適応を提案し、実験により従来の1段階のモデル適応よりも有効であることを確認した(Table 1)。

Method	w/o adaptation	1-step adaptation	2-step adaptation
Source model	Scratch	SI non-dysarthria model	SI dysarthria model
DYS-SPK1	52.35	40.20	36.41
DYS-SPK2	73.83	65.21	54.66
DYS-SPK3	80.28	75.12	67.30
DYS-SPK4	67.58	59.98	56.28

Table 1: 構音障害者4名の音声に対する、適応無し、1段階適応、2段階適応それぞれの音声認識単語誤り率[%]

### 1-P-14

#### 1-P-14 小型デバイス向け音声認識システムのための音響モデルの検討

Study on Acoustic Models for Low-Resource Device,

© 森谷崇史, 佐藤宏, 篠原雄介

日本電信電話株式会社 NTTメディアインテリジェンス研究所,

- ◆計算処理の乏しいローカルデバイス上では大規模なDNN音響モデルの動作が困難なため、コンパクトなDNN音響モデルが必要
- ◆前回はKnowledge Distillation (KD)の損失関数に中間層の出力のmean squared error (MSE)による損失を加えたMSE-KDを提案し、正解ラベルのみやKDを用いて学習したモデルよりも改善したことを報告
- ◆本研究では、提案手法(MSE-KD)により構築したDNN音響モデルにパラメタ行列の特異値分解、デコード時のパラメタの量子化を適用することで認識性能を維持しながらローカルデバイス上における認識速度の削減率72.6%を達成

Table 1: Character error rate (CER) of evaluation data.

Model	Loss	CER [%]	RTF
CNN(w/o BNL)	CE	15.6	3.878
CNN(w/ BNL)	CE	15.5	3.290
512 × 4	CE	17.6	0.351
	KD	16.7	
	MSE-KD	<b>16.1</b>	
+SVD(k=128)	MSE-KD	<b>16.1</b>	0.159
+Quantization	MSE-KD	16.2	<b>0.096</b>

### 1-P-16

#### 1-P-16 CTCとAttentionを用いたEnd-to-end ストリーミング音声認識

E2E Streaming Speech Recognition Using CTC and Attention

☆陳家浩(徳島大学), 西村良太(徳島大学), 北岡教英(豊橋技術科学大学)

- ◆本稿では、単方向LSTMとCTCを用いた局所アテンションベースのモデルを使用した日本語ストリーミング認識を提案する。
- ◆フレームスタッキングの代わりにmaxpoolingを使用するのは効果的である。
- ◆単語の頻度の不均一によって生じるオーバーフィッティングを防ぐために、信頼ペナルティを導入する。
- ◆局所アテンションを導入し、コンテキスト情報を活用することによって、低フレームレートでの音声認識性能が向上することが分かった。

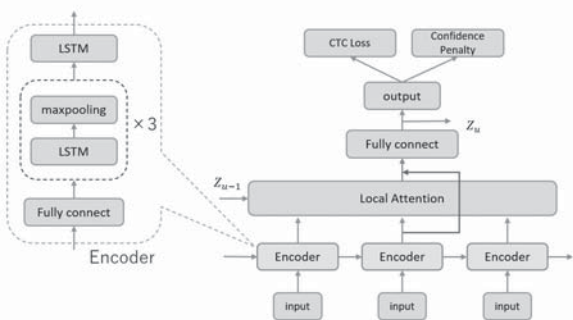


Fig. 1: Model Structure Using CTC, Confidence Penalty, and Local attention

### 1-P-15

#### 1-P-15 後向きLSTM言語モデルとアンサンブル符号化器を用いたNベストリスコアリングのための対仮説比較モデルの改良

Improved deep duel model for rescoreing N-best speech recognition list using backward LSTMLM and ensemble encoders

○小川厚徳, マーク・デルクロア, 苅田茂樹, 中谷智広

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

- ◆音声認識仮説のNベストリスコアリングを行うためのモデルとして、一対仮説比較モデル(deep duel model: DDM)を提案した。
- ◆本稿ではDDMの特徴量として、前向きLSTM言語モデルのスコアに加えて後向きLSTM言語モデルのスコアを追加する。更にDDMの構造として、複数の符号化器を並べたアンサンブル符号化器を導入する。
- ◆CSJ音声コーパスを用いた実験において、これら二つの相補的な改良により、DDMの精度向上が可能であることを示す。

Table 1. 100-best rescoreing results in WER [%]

No.	Model	Dev	Eval
1.	ASR 1-best (trigram)	18.1	14.8
2.	Forward LSTMLM	17.4	14.2
3.	Backward LSTMLM	17.3	14.1
4.	Fwd & bwd LSTMLM	17.3	14.1
5.	Single-encoder DDM	16.4	13.7
6.	5 with fwd LSTMLM score	15.6	13.0
7.	5 with bwd LSTMLM score	15.7	13.0
8.	5 with fwd & bwd LSTMLM scores	15.4	12.8
9.	Two-encoder DDM	16.2	13.5
10.	Four-encoder DDM	16.1	13.5
11.	Eight-encoder DDM	16.1	13.4
12.	11 with fwd & bwd LSTMLM scores	<b>15.2</b>	<b>12.6</b>
13.	Oracle	11.6	9.7

### 1-P-17

#### 1-P-17 高齢者音声認識のための音声コーパス構築と方言への適応の効果

A New Corpus of Elderly Japanese Speech for Acoustic Modeling, and a Preliminary Investigation of Dialect-Dependent Speech Recognition

○福田芽衣子, 西村良太(徳島大), 西崎博光(山梨大),

北岡教英(豊橋技科大), 入部百絵(愛知県立大)

- ◆高齢者音声認識精度の向上のためにS-JNAS(平均年齢:67.6歳)より更に高齢のコーパスが必要と考え、現在までに全国4カ所、平均年齢79.2歳の211名の音声をデータベース化した。
- ◆学習用データはATR音素バランス文、テストデータはJNAS収載の新聞記事文を用い、書き起こしテキストデータ3種(平仮名、カタカナおよび漢字かな混り文)を作成した。
- ◆また、JNAS、S-JNASならびにCSJを音響モデル作成に用い、これを本コーパスで適応化したところ、適応化以前はS-JNAS音響モデルの認識精度が最も良い結果であったが、本コーパスへの適応後はCSJ音響モデルが最も良い結果になった(Fig.1)。

- ◆さらに録音地域別の音声で方言への適応化の効果のみたところ、CSJ

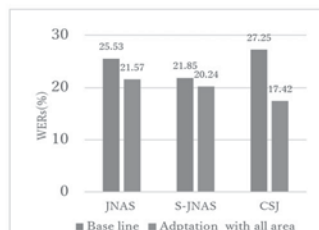


Fig. 1 WERs (%) for speech recognition experiments using JNAS, S-JNAS or CSJ trained acoustic models and BCCWJ language model with acoustic adaptations

音響モデルの認識精度がもっとも良い傾向を示した。

1-P-18

1-P-18 主観的話し者間類似度に基づく  
DNN 話し者埋め込みを用いた  
多数話し者 DNN 音声合成の実験的評価  
Evaluation of DNN-based multi-speaker speech synthesis using DNN-based  
speaker embedding considering subjective inter-speaker similarity  
©齋藤 佑樹, 高道 慎之介, 猿渡 洋(東大院・情報理工)

- ◆先行研究: 主観的話し者間類似度に基づく DNN 話し者埋め込み
  - 話し者間類似度に関する大規模主観スコアリングの結果を利用
  - 類似度スコアと対応付けられた話し者埋め込みを学習

◆本発表: この話し者埋め込みを導入した多数話し者音声生成モデルの評価

Q1: 主観的話し者間類似度の考慮で, 合成音声の音質を改善できるか?

話し者認識 に基づく学習  
VS.  
主観的話し者間類似度 に基づく学習

Q2: 主観的話し者間類似度に基づく学習に適したカーネル関数とは?

Sigmoid カーネル を用いた学習  
VS.  
Gauss カーネル を用いた学習

1-P-20

1-P-20 歌声の母音変化を考慮した  
歌声合成の検討  
Singing Voice Synthesis Considering Vowel Variations  
☆片平健太(神戸大), 足立優司(メック株式会社),  
田井清登(メック株式会社), 高島遼一(神戸大), 滝口哲也(神戸大)

- ◆女性の日本語オペラ歌唱における母音発音時のメルケプストラムに  
関して, PCA, k-means を用いて母音の変化の傾向を分析する.
- ◆クラスタリングの結果を DNN の学習時の特徴量に追加し, 母音の変  
化を考慮した DNN 歌声合成手法を検討する.

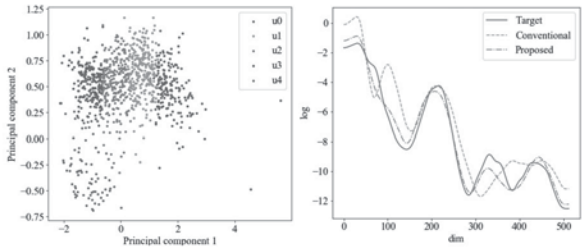


Fig.1: Clustering result of "u". Fig.2: Synthesized spectrums.

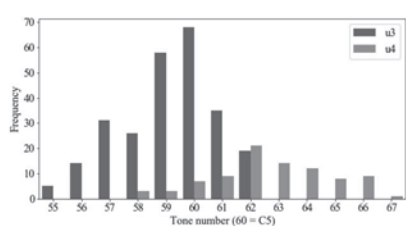


Fig.3: Frequency distribution of u3 and u4.

1-P-19

1-P-19 多様な環境音の合成と変換のための  
基礎検討  
Fundamental Study of Synthesis of Acoustic Scenes and Events  
☆岡本悠希, 柳生拓巳, 井本桂右(立命館大学)  
小松達也(LINE Corporation), 高道慎之介(東大院・情報理工)  
山西良典, 山下洋一(立命館大学)

- ◆統計的な音の合成・変換に関して, 音声を対象とした技術は幅広く検  
討が進んでいる
- ◆本稿では音声以外の多様な環境音を合成・変換する技術について整理  
➢ アプリケーション: 映画や動画などの背景音作成, 環境音データベースの拡張  
➢ 問題設定: 音響イベント/シーンラベルを用いた環境音合成, オノマトペを  
入力とした環境音合成など
- ◆Conditional WaveNet を用いた音響イベント合成を例に, 環境音合成  
技術の現状の到達点と主観評価方法についても検討
- ◆主観評価実験では音声合成で行われる評価項目を参考に, 合成された  
音響イベントに関して以下の3つの項目について評価  
➢ 合成された環境音の了解性に関する評価  
➢ 実在音としての自然性の評価  
➢ 環境音としての自然性の評価
- ◆実在音としての自然性の評価において, 82.71%の割合で本物の音と  
合成音を識別することができた  
➢ 現状の技術では本物の音と区別できないほどの高い品質の音を合成するに  
は至っていないことを確認

1-P-21

1-P-21 構音障害者の少量データを用いた深層学  
習による音声合成の検討  
Speech synthesis system using small amounts of data for Articulation  
Disorders  
☆南阪竜(神戸大), 高島遼一, 滝口哲也, 有木康雄

脳性麻痺による構音障害者の発話, 筋肉の不随意運動により喉や口  
の筋肉を正常に動かすことが出来ないため, 健常者と比べて不安定であ  
り, コミュニケーションを難しくしている. また, 多量の発話の収録は  
構音障害者に大きな負担を掛ける場合がある. そこで本研究では, 脳性  
麻痺による構音障害者を対象としたコミュニケーション支援として構  
音障害者の少量データを用いた, テキスト音声合成(TTS) システムを提  
案する.

まず健常者で音声合成モデルを生成し, そのモデルを構音障害者の少  
量データで再学習することによって, 健常者と構音障害者の2つの音声  
合成モデルを作成する. また2つのモデルによるテキスト音声合成によ  
り, 健常者と構音障害者, それぞれの正確なラベルが得られるため, こ  
れを用い音素の置き換え・基本周波数の変換を行い, 明瞭性改善を行う.

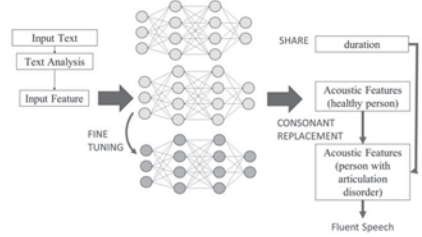


Fig1: A flow of proposal speech synthesis system

### 1-P-22

#### 1-P-22 Fader Networks を用いた楽器音変換

Instrumental sound conversion using Fader Networks

☆荒川賢也, 岸田拓也, 中鹿亘(電通大)

- ◆本研究では楽器音の性質は倍音の混ざり方や音の時間的な変化によって表現されるという特徴に着目し、スペクトログラムに対してFader Networksを適用することで、音高情報などの高次元特徴量は保持したまま、楽器の特徴を示す低次元特徴量のみを変換し、高精度な楽器ドメイン変換を行う手法を提案する。
- ◆画像のスタイル変換ではターゲットのデータが存在しなかったため変換に関する客観評価がなかったが、本研究ではMIDIデータを用いてターゲットを生成することでFader Networksの低次元特徴の変換精度に関する客観評価を行なった。
- ◆Fader Networksを用いることによって音楽音響信号の楽器音変換において従来手法(cVAE)と比較して高い精度の結果が得られた。

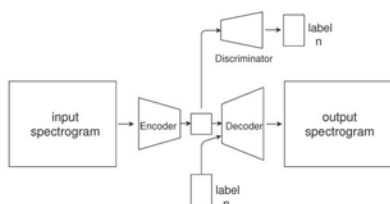


Fig.1 : The architecture of Fader Networks

Table 1 : experimental result (conversion error)

Method	RMSE
cVAE	130.43
Fader Networks	126.68

### 1-P-24

#### 1-P-24 非負値行列因子分解のアクティベーションに着目した DNN 音声合成の検討

DNN-based statistical parametric speech synthesis incorporating non-negative matrix factorization

☆後藤駿介, 齋藤大輔, 峯松信明(東大)

- ◆統計的パラメトリック音声合成においては、音響特徴量としてメルケプストラムがよく用いられる。しかし、メルケプストラムはスペクトル包絡を sin カーブの重ね合わせで効率的に表現できるが、より微細なスペクトル包絡の表現として代替の手段が期待できる。そこで本稿では非負値行列因子分解 (Non-Negative Matrix Factorization; NMF) に着目する。NMFにより得られるスペクトル基底は微細なスペクトル構造を保持しており、その重みづけであるアクティベーションを音響特徴量とした DNN 音声合成を提案する。それに加えて、提案手法において帯域拡張 (16kHz から 48kHz) の実現が可能であることを示す。

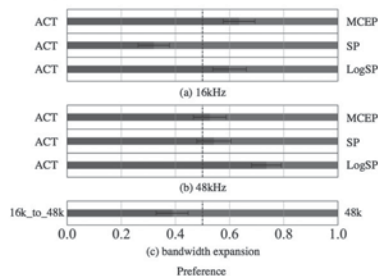


Fig. 1 Subjective results; (a) in 16 kHz, (b) in 48 kHz sampling rates and (c) of the experiment of bandwidth expansion.

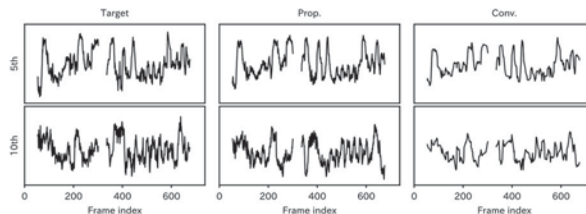
### 1-P-23

#### 1-P-23 深層学習に基づく音声合成における 2次統計量を用いたスペクトル特徴量のモデリングの検討

Study on modeling of spectral feature using second statistical features for speech synthesis based on deep learning

◎松永悟行, 大谷大和(株)エーアイ, 平原達也(富山県立大)

深層学習に基づく音声合成において、波形を直接のモデル化することにより自然音声と同等の品質の音声合成が可能になった。しかし、計算量の問題によりボコーダを用いた音声生成を利用することは少ない。スペクトル特徴量のモデル化においては、敵対的学習や畳み込み層を用いた学習方法が一般的に用いられる。これらの学習方法ではモデルの構造やパラメータ数などを調整するが、それらが出力にどのように作用するかは不明瞭であり調整が困難である。そこで、時系列であるスペクトル特徴量の隣接するフレーム間の関係性や次元間関係性を考慮してそれらを直接学習するために、生成されたスペクトル特徴量から2次統計量を計算してその誤差を最小化する学習方法を提案した。その結果、平均二乗誤差による学習方法よりも自然音声に迫るスペクトル特徴量を生成することができた。



Examples of 5th and 10th mel-cepstrum trajectories.

### 1-P-25

#### 1-P-25 深層ガウス過程に基づく音声合成におけるリカレント構造を用いた系列モデリングの検討

A study of sequential modeling using recurrent structure for deep-Gaussian-process-based speech synthesis

○郡山知樹, 猿渡 洋(東大院・情報理工)

- ▶ 深層ガウス過程(DGP)に基づく音声合成において、従来のフレーム毎の予測だけでなく、発話全体を予測できるように、時系列モデリングへの拡張を行う。
- ▶ 提案手法では、並列計算が可能で高速なSimple recurrent unit (SRU)をDGP音声合成に導入する。
- ▶ SRUにおいても、従来のDGPと同様にガウス過程回帰による分布の予測とサンプリングで順伝播を実現可能である。
- ▶ 主観評価実験によって、提案法は従来のフィードフォワード型のDGP音声合成だけでなく、LSTM-RNNやSRU-RNNよりも自然性の高い音声合成できることを示した。

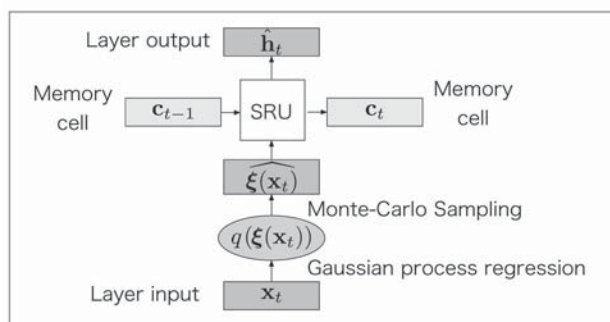


Fig. 1 Forward propagation in SRU-DGP.



## 1-P-26

### 1-P-26 ヴィブラート推定における要素パラメータの統計的な性質分析

Statistical analysis of parameters used in vibrato section estimation.

☆宮崎嵩大(山梨大), 森勢将雅(明治大)

- ◆ヴィブラート区間推定における要素パラメータとして、ヴィブラートらしさ  $P_v$  とそれを構成するヴィブラート速度の周波数帯域のパワー  $\Psi_v$ 、ピーク  $S_v$  の3つの要素パラメータが存在する。本研究はこれらの要素パラメータの非ヴィブラート・ヴィブラート音声の統計的な差を分析した。
- ◆全要素パラメータでヴィブラート・非ヴィブラートで統計的差があることから、ヴィブラートの判別に有効なことを再確認した。
- ◆今回の実験において、 $P_v$  の最適な閾値が0.1097であることを示した。

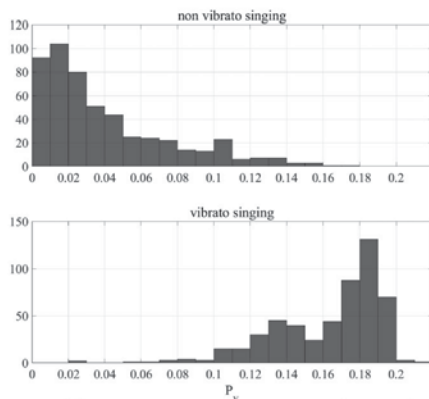


Fig. 1: Histogram of  $P_v$  (Top: non Vibrato singing, Bottom: Vibrato singing)

## 1-P-28

### 1-P-28 歌声を聴きながらヴィブラートをデザインできるインタフェースの提案

Proposal of singing design interface capable of controlling vibrato during reproduction.

☆小野雄大(山梨大), 森勢将雅(明治大)

- ◆ VOCALOID 等の歌声合成ソフトウェアの発売から、歌声合成技術は多くのコンテンツ制作に利用されている。
- ◆本研究では、歌声合成ソフトウェアにおける歌声のデザインを支援することを目的としたインタフェース開発を目指す。
- ◆前報の歌声を聴きながらヴィブラートをデザインできるインタフェース Parrot に対して、改良を行った。
- ◆Parrot と VOCALOID 4 ビブラートデザインエディタを併用する手法を提案し、Parrot V2 (Fig. 1) を開発した。
- ◆Parrot V2 は、Parrot で歌声を聴きながらヴィブラートデザインを行い、VOCALOID 4 ビブラートエディタで仕上げるといった使い方により、歌声のデザインの省力化が期待される。

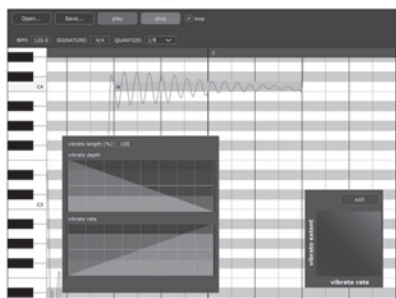


Fig. 1 The Screenshot of Parrot V2

## 1-P-27

### 1-P-27 歌唱音声におけるヴィブラートの自然性評価のためのスペクトル変動消失音声の作成方法に関する検討

Study on how to generate singing voice sound without spectral fluctuation to evaluate naturalness of vibrato in singing voice

☆鈴木千文, 坂野秀樹, 旭 健作 (名城大院), 森勢将雅 (明治大)

- ◆本研究では以下のヴィブラート合成音の自然性の評価を行った
  - スペクトル変動を含むヴィブラート合成音
  - スペクトル変動が一定であるヴィブラート合成音(スペクトル変動消失音声)
  - スペクトル変動を含むヴィブラート合成音の方が自然性が高い
- ◆そこでスペクトル変動消失音声の作成方法の検討を行った
- ◆歌声データ: ヴィブラート音声(目標歌手: Gackt)
- ◆スペクトル変動消失音声の作成
  - 基本周波数のピークを抽出し、そのフレームのスペクトルの値でヴィブラート開始以降のフレームの値を置き換える
- ◆メルケプストラム歪みとスペクトル包絡を算出(Fig.1)
  - 算出した歪み: 一定にした際に用いたフレームのメルケプストラムと各フレームのメルケプストラムの歪み メルケプストラム次数: 16 次
- ◆Fig.1(左)の方が Fig.2(右)に比べ、500[Hz]や 1500[Hz]付近のフォルマントのピークが顕著に現れている
  - このピークの有無がヴィブラートの自然性に影響する可能性あり

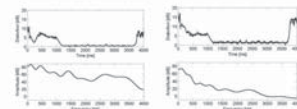


Fig.1 Mel-cepstrum distortion between the reference frame and each frame and spectral envelope of the reference frame (Left: the reference frame is the vibrato-beginning frame, Right: the reference frame is the vibrato-end frame).

## 1-P-29

### 1-P-29 文節構造を考慮した基本周波数軌跡のモデル化と焦点制御のための統計的韻律変換への応用

Generative Modeling of  $F_0$  Contours Considering Bunsetsu Information and Its Application to Statistical Focus Control

☆白旗悠真, 齋藤大輔, 峯松信明(東大)

- ◆人間の声の基本周波数 ( $F_0$ ) は、言語情報のみならず発話の態度や意図、話者性などのパラ言語情報や非言語情報を豊富に含んでおり、これらの情報は語や節、文などの大域的な時間構造に現れる。藤崎モデルは節や語をフレーズ指令とアクセント指令というパラメータに明示的に対応させたモデルであり、このモデルでは大域的な特徴を捉えるだけでなくパラメータの大きさを変化させることで言語情報に対応した柔軟な韻律変換が可能である。本稿では藤崎モデルパラメータを言語情報なしで統計的に求める手法にフレーズ指令と文節の明示的な対応を組み込んだモデルを導入し、このモデルを韻律変換に応用し文節ごとの焦点制御を行う手法について提案する。
- ◆実験による主観的評価により提案手法は自然音声に対して 74%の焦点位置正解率を達成し、提案手法による焦点制御の有効性を示した。

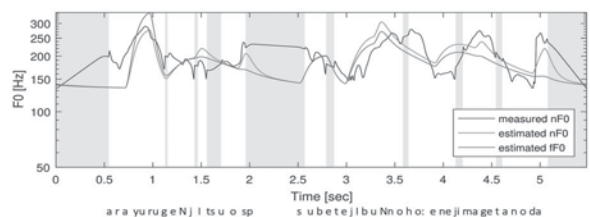


Fig 1. 提案手法によって第一文節に焦点を付与した  $F_0$  軌跡の一例

### 1-P-30

#### 1-P-30 Speech-to-Speech Translation using Dual Learning and Prosody Conversion

☆Zhaojie Luo, Yoichi Takashima, Tetsuya Takiguchi, Yasuo Ariki (Kobe University)

This paper presents an approach for speech-to-speech translation while keeping the voice sound unchanged. As shown in the Fig. 1, we combined ASR, NMT, and TTS using the dual learning method. For each two of the models, we can regard them as a pair of dual tasks. After the speech-to-speech translation, we do the voice conversion for the TTS synthesis voice. In this voice conversion task, the TTS synthesis voice is the source voice and the voice of speaker is the target. We only convert the spectral features of the synthesis voice and keep the standard pronunciation of the TTS voice unchanged. By this way, the translation voice is similar to the voice of speaker but keep the pronunciation more standard than the speaker.

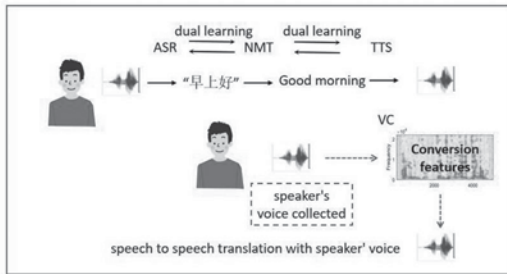


Fig. 1: Framework of the proposed speech-to-speech translation model

### 1-P-32

#### 1-P-32 DNN 音声合成を用いたアナウンサーと素人話者間の基本周波数についての分析

Analysis of fundamental frequency between a newscaster and amateur speakers using partial-substituted synthetic speech

☆大鶴拓哉(東大), 井島勇祐(NTT), 齋藤大輔, 峯松信明(東大)

- ◆DNN を用いた合成音を用いた大規模な知覚実験を行い、プロアナウンサー合成音とアマチュア話者合成音間の聞き手の印象について評価を行った。
- ◆主観評価実験の結果 (Fig.1) は、プロアナウンサーの F<sub>0</sub> パターンを合成に使用することが聴取者の印象に最も良い影響を与えたことを示した。
- ◆音響的特徴量と聴取者の主観スコアとの相関を明らかにするため、実験に用いた合成音を分析した。その結果、F<sub>0</sub> の標準偏差と音素単位のリズムが主観スコアと高い相関を示すことが分かった。

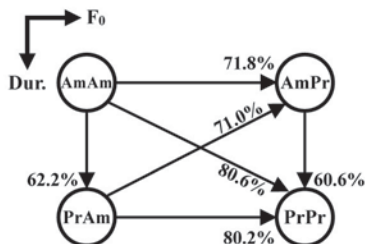


Fig. 1: The results of the perceptual experiment. The horizontal axis indicates substitution of F<sub>0</sub>, and the vertical one indicates substitution of phoneme duration. The direction of arrows indicates that which speech stimuli is acceptable as speech of a newscaster in each paired comparison. Scores are average values from 5 amateur speakers.

### 1-P-31

#### 1-P-31 聴覚障害高校生の文章音読における発話速度の検討

A Study of the Speech Rate of Reading Speech in High School Students with Hearing Impaired

☆湯浅哲也, 加藤靖佳(筑波大)

- ◆聴覚障害児は発話明瞭度をより高める必要性から、発話速度の調節や発話速度に対する意識を高めることが求められる。
- ◆しかし、これまでの聴覚障害児の発話速度の研究において、中学生・高校生を対象にした研究はほとんど見られない。
- ◆本研究では、聴覚特別支援学校高等部生徒を対象に文章音読を用いて、発話速度及び背景要因との関連について検討することを目的とした。
- ◆発話速度は、129~252ms/mora と個人差が大きかった。
- ◆背景要因との相関分析の結果、発話速度の自己評価と1モーラ当りの持続時間に1%水準で有意差が見られた。

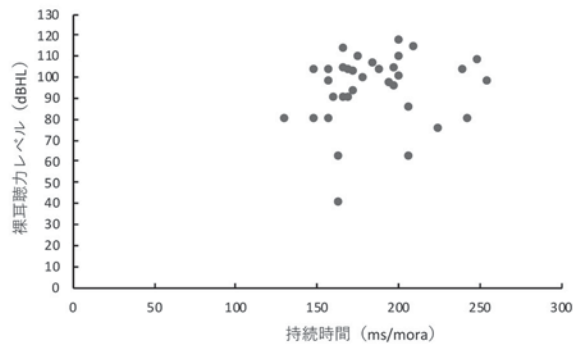


Fig.1: Speech rate

### 1-P-33

#### 自発音声に対するニューラル F0 モデリングの可能性

◎永田智洋, 森大毅 (宇都宮大)

##### 目的

音声の韻律的特徴は話者の意図や感情といったパラ言語的メッセージを伝達する重要なキャリアであり、特に自発音声の韻律的特徴の適切なモデル化は様々な音声研究において有用である。自発音声コーパスはラベリングのコスト等の理由により、限られた情報しか利用できない場合もある。そこで、本研究では簡単に利用できる情報とニューラルネットワークを利用した自発音声に対する F0 モデリングを行う。

##### 手法

句境界情報を入力として F0 を出力するニューラルネットワークを構築する。ネットワーク構造には時系列の特徴を捉えるために再帰構造を採用する (Fig. 1)。

##### 結果

句境界情報のみを用いて構築された F0 モデルは、実音声の F0 パターンを概ね表現可能であることを示した (Fig. 2)。

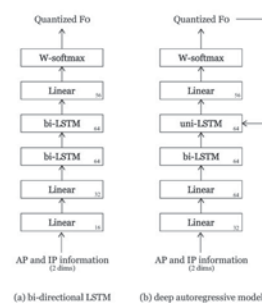


Fig. 1 The constitution of neural network.

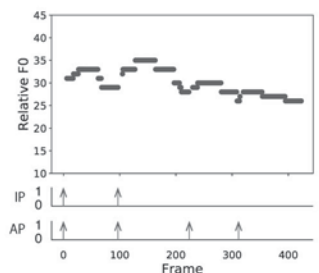


Fig. 2 The example of generated F0.

1-P-34

1-P-34 音声合成におけるブロック処理によるパラメータ軌跡生成の検討

A study for blockwise parameter generation for speech synthesis  
 ○西澤信行, △小原朋広, △服部元 (KDDI総合研究所)

HMM 音声合成に必要な RAM サイズを削減するには、従来は発話単位で行っていたパラメータ軌跡生成を、ブロック単位で行えばよい。各ブロックの計算でブロック外側の D フレームも併せてまず計算を行い、この D フレームの計算結果は使わずに捨てる方法を用いた場合、D を大きくすると従来法に対するメルケプストラムひずみ (MCD) は急激に小さくなる。処理量は従来法の 1+D/L 倍 (L: ブロック周期) なので L を大きくすれば処理量の増加を抑えることができ、実際のシステムで L を大きくすることによるその他の影響はそれほど大きくない。実用上問題のない水準の処理量と RAM サイズの増加と遅延時間で、十分に小さい MCD を達成できる。

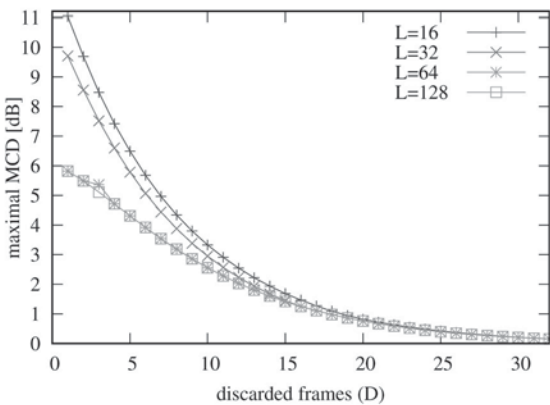


Fig. 1 Maximal MCD in synthetic sounds of 503 sentences

1-Q-2

1-Q-2 広帯域騒音におけるバーチャルセンシング技術の検討

Comparisons of Virtual Sensing Methods for Broadband Noise  
 ☆前田 怜央, 梶川嘉延 (関西大・システム理工)

- ◆騒音に対して同振幅、逆位相の擬似騒音を重ね合わせることで騒音低減を行う ANC をヘッドレストに導入し、使用者の騒音による不快感や健康被害への解決案の一つとして研究されている。
- ◆通常の ANC システムでは誤差マイクロホン地点を中心に消音空間が生成されるため、所望地点において低減効果が十分でない可能性がある。本稿では、2つのバーチャルセンシング技術 (リモートマイクロホン法、補助フィルタ法) を ANC ヘッドレストシステムに導入し比較を行った。
- ◆2つのバーチャルセンシング技術を用いた ANC ヘッドレストシステムの広帯域騒音に対する有効性について MATLAB でシミュレーションを行った。その結果、Fig.1 に示されるように所望地点において補助フィルタ法の方がリモートマイクロホン法より騒音源の位置にかかわらずより大きな低減効果が得られることがわかった。

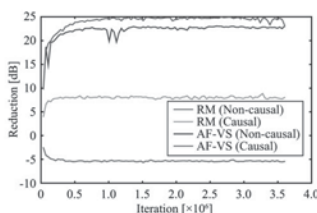


Fig.1: Comparison of the reduction level at the right virtual microphone when the direction of a user's head is forward.

1-Q-1

1-Q-1 気導音と骨導音に基づくベイズ定理に UT 法を融合した雑音除去法

Noise cancellation method joining Unscented Transformation Method and Bayes Theorem based on air- and bone-conducted speech signals.

○折本寿子, 生田顯 (県立広島大)

騒音下の音環境における音声認識には、周囲の環境雑音対策は不可欠であり、気導音だけでなく外来雑音の影響が少ない骨導音を併用し、UT 法をベイズ定理の展開表現に融合させることにより高次の相関情報が含まれる新たな雑音除去を提案した。

理論の有効性を確認するため、女性、男性の音声信号に白色雑音、有色雑音、機械雑音を混入させることにより実験を行っている。

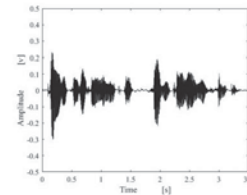


Fig.1 Original male speech signal.

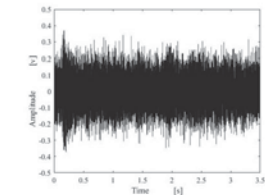


Fig.2 Observed male speech signal contaminated by a white noise with amplitude of 3 times.

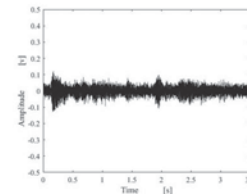


Fig.3 Estimated results by use of proposed method.

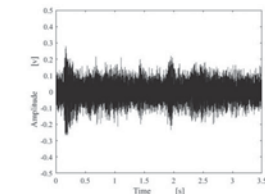


Fig.4 Estimated results by use of compared method.

1-Q-3

1-Q-3 適応型インパルス応答測定信号における雑音スペクトル推定方法の検討

A study of noise spectrum estimation for adaptive impulse response measurement signal

☆大森一弘, 金田豊 (東京電機大・工)

- ◆音響系のインパルス応答を測定する際、環境雑音のスペクトルをあらかじめ推定し、それを基に測定信号のスペクトルを決定することで雑音成分を最小化する MN-SS (Minimum Noise - Sweep Sine) 信号が存在する。
- ◆雑音スペクトルの推定方法および推定次数の違いによって推定した雑音スペクトルの形状は Fig. 1 のように大きく異なる。そして、それを用いたインパルス応答測定結果の雑音低減量もそれぞれで異なってしまう。
- ◆複数の騒音に対して、それぞれの方法で測定信号を作成し、測定したインパルス応答の雑音成分のパワーは Fig. 2 のようになり、図より線形予測法が自己相関法より優れていることがわかる。
- ◆また、雑音成分のパワーに加え、測定信号の掃引正弦波としての特性、周期雑音成分の低減を考慮すると雑音スペクトル推定方法は、推定次数 300 程度の線形予測法を用いることが適切である。

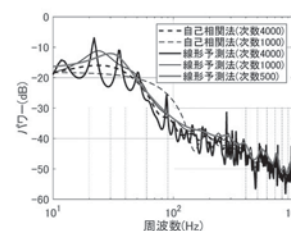


Fig. 1 推定方法による雑音スペクトルの違い

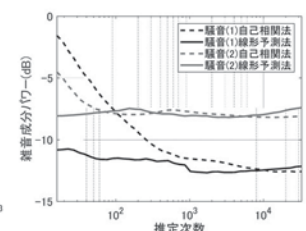


Fig. 2 各方法の推定次数と雑音成分パワー

### 1-Q-4

#### 1-Q-4 エンドファイア型スピーカアレイを用いた 能動騒音制御の検討

Study on Active Noise Control Using Endfire Loudspeaker Array

©秋元亜門, 羽田陽一(電通大)

- ◆本研究では、発話音声は既知であり、かつ頭部を剛球とみなしたときにスピーカアレイを2次音源として用いた手法で音声抑圧が可能なのかを検討する。
- ◆剛球から放射された信号の音圧を、後方にある制御点の位置で小さくするため、周波数領域での最小二乗法を用いてフィルタを導出した。
- ◆実験は無響室で行い、剛球モデルとエンドファイア型スピーカアレイを設置し、剛球周辺での音圧を円形マイクアレイを用いて測定した。
- ◆実験結果として、図2のように抑圧方向(270°方向)での音圧レベルの減衰が確認できた。抑圧方向以外では、2次スピーカ数が少ないために、音圧レベルが高くなる傾向が見られた。

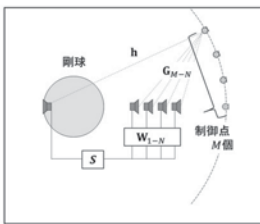


Fig.1: System Diagram

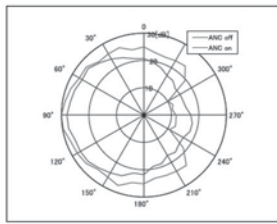


Fig.2: Sound Pressure differences by ANC on/off. (Blue : ANC off, Orange : ANC on)

### 1-Q-6

#### 1-Q-6 周波数を冗長化したSTFTによる 位相復元の音声強調への利用

Application of phase reconstruction with redundant short-time Fourier transform along frequency axis to speech enhancement

©若林佑幸, 小野順貴(首都大)

- ◆我々は前回の発表において、零詰め処理と冗長度パラメータ $\alpha$ を導入し、周波数方向の冗長度を考慮した短時間フーリエ変換(Short-time Fourier transform: STFT)による時間周波数表現を提案した。
- ◆STFTの位相に着目した研究において、分析フレームシフトはフレーム長の1/4, 1/8の長さで設定される。時間方向への冗長度を大きくしていると解釈でき、これにより位相の時間方向の動きを詳細に分析することができる。提案するSTFTは位相の周波数方向の詳細な分析を行うことができると期待される。
- ◆本研究では周波数冗長化STFTの応用として音声強調を取り上げ、既存の振幅推定手法とGriffin-Limアルゴリズムと併せて、処理を行った。
- ◆雑音重畳音声に対し強調処理を行ったところ、Fig. 1のように通常のSTFT( $\alpha=1$ )や観測位相(実線、鎖線、破線はそれぞれ $\alpha=1, 2, 4$ のときのもの)を用いたときと比較して周波数冗長化STFT( $\alpha=2, 4$ )によってPESQが改善した。

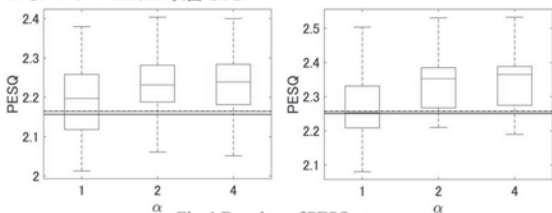


Fig. 1 Boxplots of PESQ score.

### 1-Q-5

#### 1-Q-5 音源信号の事前知識に基づく多重制約付き 最適化によるPSD推定

Power spectral density estimation by convex optimization with multiple constraints based on prior knowledge of source signals

©千葉大将, 丹羽健太, 原田登 (NTT), Guoqiang Zhang (シドニー工科大), W. Bastiaan Kleijn (ヴィクトリア大学ウエルントン校)

- ◆目的音到来方向が既知の条件でのリアルタイム音源強調手法として、ビームフォーミングとウィナーポストフィルタを併用した手法がある。この手法の音源強調性能は、目的音、干渉雑音、および背景雑音のPSD(パワースペクトル密度)の推定精度に依存している。
- ◆本研究では、音声信号(目的音)や雑音、收音環境に関する様々な事前知識を使って、高精度にPSDを推定する手法を提案する。コスト関数は線形制約付き凸最小化問題に帰着され、その双対問題を高収束率であるプレグマン単調作用素分解を用いて最適化した。
- ◆一般的な音声信号で仮定できそうないつかの事前知識を凸コスト/線形制約としてPSDを推定した結果、PSDの推定精度および音源強調性能が向上していることを確認した。

Table 2 SDR improvement scores [dB]

	RT <sub>60</sub> : 110 ms		RT <sub>60</sub> : 330 ms	
	K=3	K=5	K=3	K=5
(Obs. SDR)	-5.97	-7.69	-5.79	-7.65
BF	5.69	6.34	4.06	4.70
Conv.	9.57	10.20	8.39	9.14
Proposed	9.96	10.88	8.74	9.71

Table 3 PSD distortion scores [dB]

	RT <sub>60</sub> : 110 ms		RT <sub>60</sub> : 330 ms	
	K=3	K=5	K=3	K=5
Conv.	8.72	7.35	7.98	6.96
Proposed	10.49	9.58	9.44	9.11

### 1-Q-7

#### 1-Q-7 光レーザーマイクロホンを用いた深層学習と 調波構造復元による音声強調

Speech enhancement based on deep learning and harmonic reconstruction using optical laser microphone

☆上田 奨士(立命館大院), 福森 隆寛, 西浦 敬信(立命館大)

- 光レーザーマイクロホンは対象物体にレーザー光を照射し、その微小な振動を計測することで音を収録するマイクロホンであり、遠方音声の計測に有効である。しかし、その受音音声は振動物体の材質や計測時の環境によりFig. 1のような音質劣化が発生する。音質を改善する手法として、深層ニューラルネットワーク(DNN)を用いた音声強調が提案されているが、光レーザーマイクロホンによる受音音声は高域が欠落するため、高域復元が不十分となるという問題がある。この問題を解決するために、本稿ではDNN適用後に調波構造復元を加えた新たな音声強調手法を提案した。提案法の有効性を示すために、Fig. 1のような音質劣化が発生する状況で光レーザーマイクロホンにより収録し、その音声にDNN提案法を適用し、Wideband PESQによる評価を実施した。Fig. 2より、対数パワースペクトルを用いた学習では、(a) - (e)の5種類の音質劣化において(a), (b), (c), (e)で音質を改善できていることから提案法の有効性を確認した。

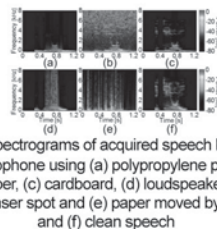


Fig. 1: Spectrograms of acquired speech by laser microphone using (a) polypropylene plate, (b) paper, (c) cardboard, (d) loudspeaker with large laser spot and (e) paper moved by wind, and (f) clean speech

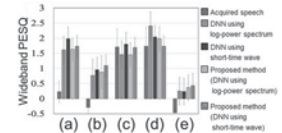


Fig. 2: Wideband PESQ scores of acquired speech, speech applied DNN and proposed method

1-Q-8

1-Q-8 Modified-VQVAE を用いた辞書学習に基づく  
音声強調手法

Modified VQ-VAE for Monaural Speech Enhancement based on Dictionary Learning

☆村井健太郎, 竹内太法, 立蔵洋介(静岡大)

- ◆研究概要
 

本研究では VQ-VAE を利用し、音声特徴量の辞書に基づいて音声強調を行う新たな枠組みを提案する。
- ◆提案手法
  - > 事前学習
 

特徴抽出モデルを利用し、60 パターンの音素ごとに 100 本の特徴ベクトルを生成し、特徴量辞書を構築する。
  - > 本学習
 

Modified VQ-VAE を利用した辞書学習による音声強調を行う。
- ◆実験評価
 

TIMIT データセットを用いて、性能を評価し、高い雑音除去性能を確認した。

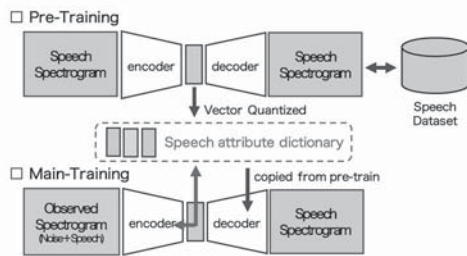


Fig.1: Overview of Proposed method

1-Q-10

1-Q-10 音源移動条件下での非同期分散マイクロホンアレイの同期処理とそれに基づく音源分離

Speech separation by synchronizing distributed asynchronous microphones under existence of source movements

○荒木章子, 小野順貴\*, 木下慶介, デルクロアマーク\*

\* 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  
\*\* 首都大学東京

- ◆非同期マイクロホンの同期をとるための技術として、筆者らは以前、収録中に音源の移動がある場合でも頑健に端末間のサンプリング周波数ミスマッチを推定可能な方法を提案した。
- ◆本稿では、このミスマッチ量の妥当性の評価の一環として、当該手法で推定したミスマッチ量を用いて端末間同期をとることで、端末間同期をとらない場合よりも、アレイ処理(本稿では音源分離処理)の性能を向上できることを報告する。

Table: Separation performance with and without synchronization

devices	metric	input	w/o sync.	w. sync.
Reference	SDR [dB]	0.000	2.721	6.900
IC & IC ①	SIR [dB]	-0.004	6.904	15.222
Reference	SDR [dB]	0.000	3.670	7.673
IC & IC ②	SIR [dB]	-0.004	8.224	15.860

1-Q-9

MATLAB 瞬時周波数 Toolbox

MATLAB toolbox for calculating instantaneous frequency

○矢田部浩平, 升山義紀, 草野翼, 及川靖広(早大理工)

- ◆スペクトログラムの位相と瞬時周波数の解説論文
  - ▶ 日本音響学会誌, **75**(3), 147-155 (2019)
  - ▶ Acoust. Sci. & Tech., **40**(3), 170-177 (2019)
- ◆解説論文の出版に合わせて MATLAB コードを公開
  - ▶ 解説論文の内容を補うためのデモコードを含む
  - ▶ ヘルプを丁寧に書いて使いやすく工夫してある
  - ▶ しかし、説明文がすべて英語で書かれている
- ◆公開した MATLAB コードの日本語説明書を執筆
  - ▶ 本発表の講演論文として日本語の説明書を与えた
  - ▶ 収録したコサイン級数窓 19 種類を一覧で表示
  - ▶ 窓関数の微分可能性を新たに調査して記載した

1-Q-11

1-Q-11 船舶航行音による窓振動調査のためのシステム構築\*

\*System construction for investigation of window vibration by boat noise

☆深津遼貴, 武藤憲司(芝浦工大)

- ある運河では小型船舶が朝から夕方にかけて多く航行し、運河沿いの住宅に騒音影響をもたらしている。本研究の目的は船舶航行音を適応騒音制御し室内空間を静音化するシステム構築である。本報告では、窓振動調査のためのシミュレーションシステムの構築と窓振動制御を用いた単一周波数振動低減量を報告する。
- 構築した窓振動 ANC システムの模擬装置で、振動レベルと音圧レベルの低減を確認できた。また振動分布の腹と節の腹よりの位置にアクチュエータを設置することで制御が可能と考えた。

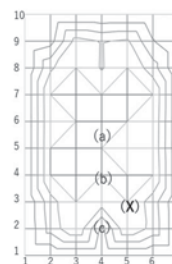


Fig. 1 Vibration distribution map by 120 Hz and Actuator's position

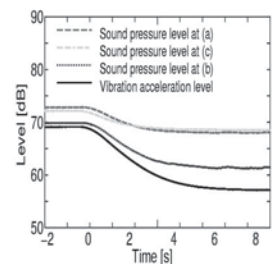


Fig. 2 Vibration acceleration level and sound level (Actuator's position :X)

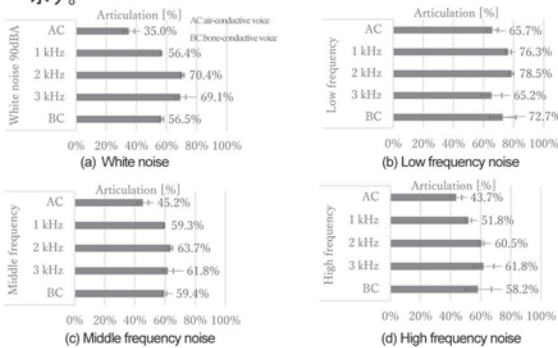
### 1-Q-12

#### 1-Q-12 環境騒音を模した騒音下での耐騒音性デュアルマイクロホンの性能の検討

Study of the performance of anti-noise dual microphone in imitated environmental noises

☆松本飛竜、青木茂明、島内末廣(金沢工大)

- ◆ 騒音下でも通信端末で快適な通話を行うために、環境騒音を模した騒音を用いて耐騒音性デュアルマイクロホンの性能を検討した。
- ◆ 無騒音下、ホワイトノイズ騒音下と 1 kHz から 4 kHz 前後にピークがある高周波数帯域の騒音、0.63 kHz から 1 kHz 前後にピークがある中周波数帯域の騒音、0.2 kHz から 0.5 kHz 前後にピークがある低周波数帯域の騒音の 5 種類について単音節明瞭度試験を行った。
- ◆ 試験時は骨導音・気導音のほかに低帯域を骨導音、高帯域を気導音とした遮断周波数 1 kHz、2 kHz、3 kHz の合成音を使用した。
- ◆ 試験結果より、騒音に含まれる周波数帯域によって明瞭度の傾向が異なることがわかった。例として騒音レベル 90 dBA の試験結果を示す。



### 1-Q-14

#### 1-Q-14 音響インパルス応答測定における測定時間短縮方法の検討

A study on measurement time reduction of room impulse response

☆田中阿子,森谷晃行,金田豊(東京電機大)

- ◆ 音響インパルス応答測定において、円状畳み込みを実現するため測定信号を 2 周期再生し、測定信号の 2 周期目を切り出していた。
- ◆ しかし、高 SN の測定を行うために測定信号長を長くする場合があります。複数周期再生すると測定時間が長くなるという問題がある。
- ◆ 本報告では、測定応答の 1 周期目と 2 周期目に同一の部分が存在することに着目した。1 周期+インパルス応答長のみ再生し、1 周期目からはみ出した部分を 1 周期目の同区間に置き換える手法を提案した。この方法により、円状畳み込みを実現できる。
- ◆ 検証実験を行った。結果を Fig. 1 に示す。提案法の 1 周期+インパルス応答長(N+Lh)再生は、従来法の 2 周期(2N)再生と比較しても同程度の誤差で測定でき、測定時間短縮に有効であることを示した。

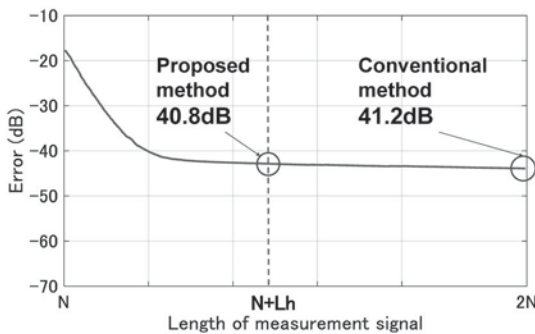


Fig. 1 測定時間長と誤差の関係

### 1-Q-13

#### 1-Q-13 反射音環境におけるスピーカ周波数特性測定手法の検討

A study on frequency characteristic measurement of a loudspeaker in a reflective environment

☆森谷晃行, 金田豊 (東京電機大)

- ◆ スピーカ周波数特性の測定には通常、無響室や無響箱を用いる。これらの設備は高価であるため、反射音環境で測定を行う手段が望まれる。しかし、反射音環境では周波数特性が劣化する(Fig.1 灰細線)。
- ◆ 反射音の影響を取り除く方法として、反射音が到達するまでの直接音を切り出す方法があるが、低域の特性が劣化する(Fig.1 破線)。
- ◆ 別の方法として、反射音到達までの直接音を使って反射音を打ち消す方法が考えられる。しかし、反射音は直接音と比較して高域のパワーが落ちるため、そのままではうまくいかない。
- ◆ そこで、反射音の影響は低域部分を中心であることに着目し、低域のみで反射音を打ち消し、直接音の高域部分と接合する手法を提案した。
- ◆ 反射音環境でインパルス応答を測定し、提案法による反射音低減を行った。Fig.2 に示した結果より、反射音到達までの直接音を切り出す場合と比較して 150~400 Hz 付近のスピーカ周波数特性を回復できることを確認した。

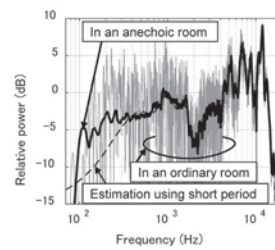


Fig.1 Frequency characteristics of a loudspeaker in an ordinary and an anechoic room.

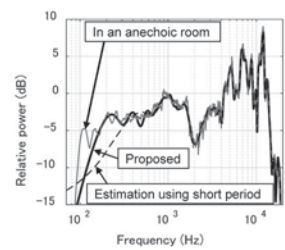


Fig.2 Improvement of frequency characteristic by the proposed method.

### 1-Q-15

#### 1-Q-15 位相の微分を用いた非混合な時間周波数ビンの検出

Detection of un-mixed time frequency bins based on phase derivative of multichannel signals

☆蛸間涼, 矢田部浩平, 及川靖広 (早大理工)

- ◆ 複数チャンネル信号処理では一般的にチャンネル間の情報を用いることが多い。ただし、チャンネル間の情報は音源が混ざっていない時間周波数ビンのものを扱うことが望ましい。
- ◆ 複数チャンネル信号を対象として位相情報を用いて音源が混ざっていない時間周波数ビンを検出する手法を提案する。
- ◆ 位相の微分情報を用いることで空間情報を無視した非混合な時間周波数ビンの検出を可能とした。
- ◆ Fig.1 に示すように音源方向推定に導入することで有効性を確認した。

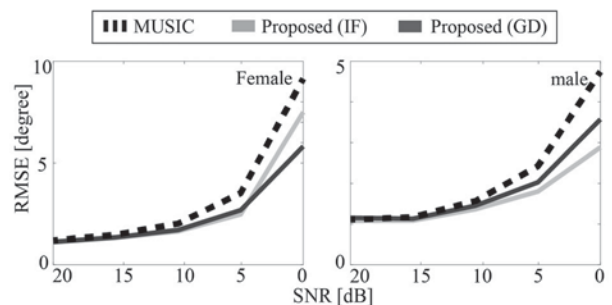


Fig.1: RMSE of estimated DOA by MUSIC method for each SNR

1-Q-16

1-Q-16 複素スペクトログラムの  
スパース・低ランクモデリング

Low-rank and sparse modeling of complex-valued spectrogram

◎ 升山義紀, 矢田部浩平, 及川靖広 (早大理工)

- ◆ 複素スペクトログラムの低ランクモデリング
  - 瞬時周波数を利用した位相修正(IPC)を利用し、複素スペクトログラムを複素行列のまま低ランクモデリング
  - 打ち切り特異値分解(TSVD)を利用した簡単な低ランク近似
- ◆ 提案手法
  - 雑音除去を位相修正した複素スペクトログラムが低ランクとなる時間領域信号を求める凸最適化問題として定式化(Method 1).
  - スパース性を誘導する正則化項を容易に追加可能(Method 2).
  - 楽器音の雑音除去で従来手法を上回る性能を確認

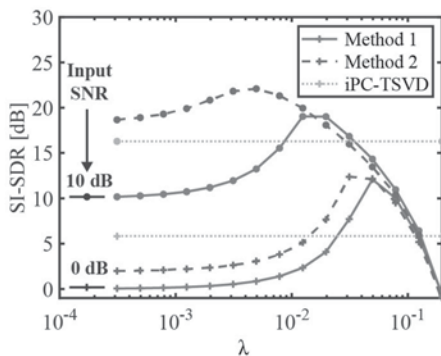


Fig. 1: Denoising results of 6 guitar sounds corrupted by additive Gaussian noise.

1-Q-18

1-Q-18 近接 3ch マイクロホンを用いた  
統計的独立性に基づく音響測距における  
測定系の影響除去の試み

Fundamental Study of Removing the Effect of Measurement System in Acoustic Distance Measurement Based on the Statistical Independency Using Adjacent 3ch Microphones.

☆ 田井 一, 篠原 寿広, 上保 徹志, 中迫 昇 (近畿大)

- ◆ 様々な工学分野において対象物までの距離情報は重要であり、送信波と反射波の位相干渉に基づく音響測距法が提案されている。
- ◆ 実環境において、雑音の混入は不可避である。雑音を低減する手法として、加算平均が挙げられるが、複数回の測定が必要となる。
- ◆ 雑音の低減手法として、源信号の統計的独立性に基づく畳み込み混合を考慮した複素 ICA(Independent Component Analysis)が知られている。
- ◆ 本研究では、高雑音下において 1 回の測定で距離推定を行うことを目標として、3ch 観測信号に 2ch の ICA を 2 度適用し、マイクロホン位置での 2 つの測距信号によって実用的に距離推定を行う手法を提案する。
- ◆ Fig. 1 に前処理なしの距離スペクトル, Fig. 2 に複素 ICA による雑音低減後の距離スペクトル, Fig. 3 に提案手法による雑音低減後、測定系の影響を除去した距離スペクトルを示す。Fig. 3 より、真値 1m 付近にピークを確認することができる。実験結果より、提案手法の有効性を確認した。

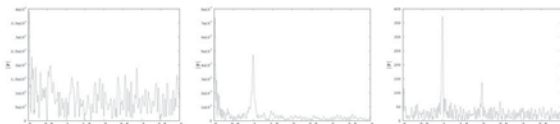


Fig. 1: Range spectrum without ICA. Fig. 2: Range spectrum with ICA. Fig. 3: Range spectrum with ICA & Cross Spectrum.

1-Q-17

1-Q-17 オンライン音声強調方式に対する反復位相復元

Iterative Spectral Phase Reconstruction for Online Speech Enhancement Methods.

○ 石橋明希, 川村新 (京都産業大)

- ◆ オンラインで実行可能な反復位相復元法である RTISI-LA を音声強調システムに導入した場合の有効性について検討する。
- ◆ 音声強調システムが理想的な音声スペクトルを推定できたと仮定し、観測信号の位相スペクトルから矛盾のない位相スペクトルを復元した場合の効果について検証した。
- ◆ 検証結果から、特に 0dB 以下の低 SNR の観測信号に対して、RTISI-LA 導入が有効であることが示された。

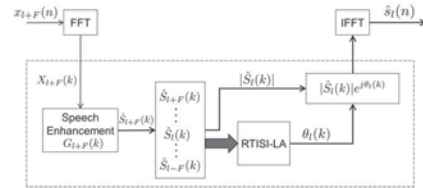


Fig. 1: Block diagram of speech enhancement system with RTISI-LA.

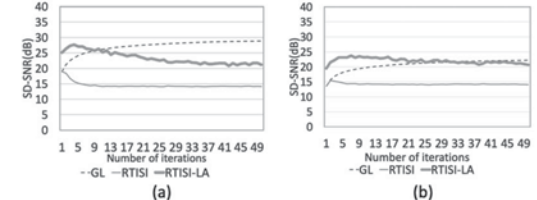


Fig. 2: Results of SD-SNR for RTISI-LA, RTISI and GL methods (a)Input SNR=0dB. (b)Input SNR=-10dB.

1-Q-19

1-Q-19 スピーカのインパルス応答長短縮の検討

A study of impulse response length compression of loudspeaker

☆ 伊藤巧, 金田豊 (東京電機大・工)

- ◆ スピーカのインパルス応答測定においては、一般室内での反射音の影響を避けるため、無響室を用いての測定が必要である。
- ◆ 本研究ではスピーカインパルス応答の第一反射音が到来するまでの直接音部分を用いて、スピーカの周波数位相特性が推定できるものと仮定し、それを用いた逆位相フィルタの有効性を検討した。
- ◆ Fig. 1 に逆位相フィルタを適用したインパルス応答波形 (シミュレーション) を示す。逆位相フィルタにより、直接音および各反射音はピークに対し対称の波形となる。直接音の前半部分を切り出し、左右反転させた波形を前半部と合成することで反射音の影響が小さいインパルス応答が得られ、これより良好な周波数振幅特性(Fig. 2)を求めることができる。
- ◆ 以上、スピーカの位相特性を推定出来れば、一般室内においてもスピーカの振幅特性を求めることが可能であることを確認した。

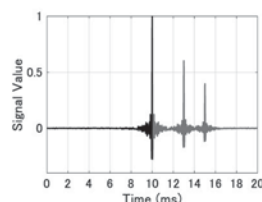


Fig. 1: 逆位相フィルタを適用したインパルス応答波形

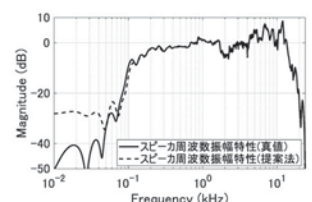


Fig. 2: スピーカの周波数振幅特性

### 1-Q-20

#### 1-Q-20 広帯域スピーカによる超音波成分も利用した音響測距

Acoustic Distance Measurement Based on the Phase Interference Using Both Audible Sound Component and Ultrasonic Component with Loudspeaker of Broadband Type.

☆北野 翔子, 篠原 寿広, 上保 徹志, 中迫 昇(近畿大)

- ◆様々な工学分野において対象物までの距離情報は重要であり, 送信波と反射波の位相干渉に基づく音響測距法が提案されている。
- ◆この手法では, 本来, 送信音は広帯域信号であり, その振幅スペクトルはすべての周波数において一様である必要がある。
- ◆距離分解能は帯域幅と逆比例の関係にある。通常のスピーカからの可聴音を使った計測を想定した場合, 利用できる帯域に制約があった。
- ◆本研究では, 送信波の帯域幅を広げ前回と同じ帯域幅で超音波領域を含む範囲にシフトした場合, 測距可能か検討する。
- ◆実験では, 可聴音のみの送信波と超音波を含む送信波を用いて測距を行った。マイクロホンと対象物の距離を 0.970~1.030[m]まで 0.005[m]刻みで変化させた。Fig. 1 に帯域範囲が2153~18653[Hz], Fig. 2 に帯域範囲が7644~24144[Hz]の場合の結果を示す(赤線:理論値, 青線:推定値, 緑線:許容誤差範囲)。実験結果より, 超音波を含む送信波を用いた測距の有効性を確認した。

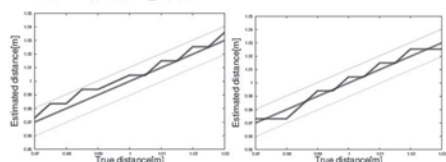


Fig.1: Comparison between true and estimated distances(2153~18653[Hz]). Fig.2: Comparison between true and estimated distances(7644~24144[Hz]).

### 1-Q-22

#### 1-Q-22 SBPH-MUSIC 法による複数音源方向推定の検討

A study of multi sound source direction estimation using SBPH-MUSIC

☆川田 希望, 金田 豊(東京電機大)

- ◆音源方向推定を行う際, 反射音の影響によって誤推定が発生する。
- ◆SBPH (Sub Band Peak Hold) 処理は, 直接音を利用して反射音を抑圧する。一方向の音源の到来方向推定で, その有効性が確認されていた。
- ◆複数方向の音源の到来方向推定の場合, SBPH 処理で一方の音の成分が他方の音の成分をマスクしてしまう可能性があった。
- ◆実録音声(近距離反射強条件)を使用して MUSIC 法と SBPH 処理を適用した SBPH-MUSIC 法で推定性能の比較を行った。
- ◆MUSIC 法で誤推定数が 280 個中 110 個あったが SBPH-MUSIC 法を使用することで 35 個に減少した。この結果から SBPH 処理は複数方向の音源方向推定においても反射音耐性に優れる処理であることが確認できた (Fig.1)。

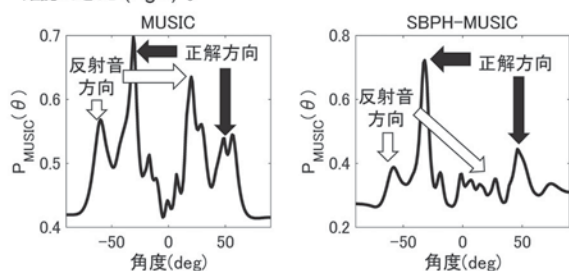


Fig.1 PH 処理適用有無の複数方向推定

### 1-Q-21

#### 1-Q-21 DNN と CAE を用いた構造物内部の音源位置推定手法に関する精度評価

Evaluation of Internal Sound Source Localization Method by Using DNN and CAE

◎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)を利用した音源探査手法による構造物内部に存在する音源の探査精度は, 90%以上を有すると, 実験と数値実験よりわかった。
- ◆これまででは, 分類モデルを用いていたため, 事前に音響部分空間を定義する必要があった。音響部分空間を定義する場合, その大きさやデータ数の違いにより, 学習可否や推定精度のばらつきが発生する。
- ◆この課題解決を目的として, 音源の座標値を直接推定する回帰モデルを構築し, 音源探査が可能であるか数値実験的に検証した。
- ◆結果としては, 過学習は発生せずに学習が可能であり (Fig. 1), 音源探査精度は約 15 mm であった (Table 1)。また, 音源探査精度の悪い箇所は構造物内部の壁面近傍に分布していることがわかった (Fig. 2)。

Table 1 RMSE & Loss at 500 epoch.

	Learning	Validation	Test
RMSE	15.12	15.22	15.24
Loss	228.64	231.89	232.45

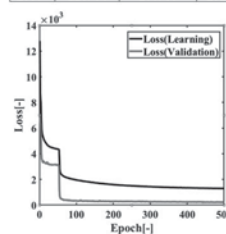


Fig.1 Learning and Validation curve.

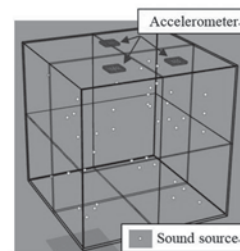


Fig.2 Sound source position with RMSE more than 50mm.

### 1-Q-23

#### 1-Q-23 ランク制約付き空間共分散モデル推定の逆行列展開による高速化

Acceleration of rank-constrained spatial covariance model estimation by expansion of matrix inversion

◎久保優騎(東大), 高宗典玄(東大), 北村大地(香川高専), 猿渡洋(東大)

- ◆1つの方向性の目的音源と, 全方位から到来するような拡散性を持つ雑音が混在する状況を扱う。
- ◆ランク制約付き空間共分散モデル推定法は, ランク 1 の目的音源空間相関行列とランクが一つ落ちた雑音空間相関行列が独立低ランク行列分析によって正確に推定できることに基づく手法である。独立低ランク行列分析の後処理として EM アルゴリズムを用いて欠けた雑音のランク 1 空間成分を補うことで, 効率的な目的音の抽出を行う。
- ◆本稿では, 各時間周波数点における逆行列を Sherman-Morrison の公式と疑似逆行列を用いて展開することで, ランク制約付き空間共分散モデル推定の時間計算量を大幅に削減し, マイク数に依存しないオーダーの計算時間でパラメータ推定を可能にする。
- ◆拡散性を持つ人混み雑音に音声信号が混在する状況を模擬した実験により, 提案手法が計算速度の面で従来手法に優れることを確認する。

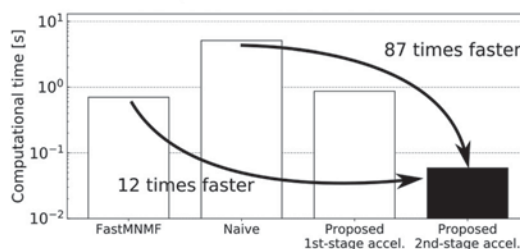


Fig. 1 Average computational time of one iteration for each method.



### 1-Q-24

#### 1-Q-24 インパルス応答のスパース性を用いた教師あり独立低ランク行列分析

Supervised independent low-rank matrix factorization using the sparsity of impulse responses

☆高橋 悠希, △北原 大地, 平林 晃(立命館大)

- ◆音源分離問題は学習データを用いない「ブラインド音源分離」と学習データを用いる「教師あり音源分離」に大別される。
- ◆ブラインド音源分離手法の独立低ランク行列分析 (ILRMA) は各音源の統計的独立性と低ランクな時間周波数構造を仮定した手法である。
- ◆上記の仮定に加え、インパルス応答のスパース性を考慮することで、ILRMA の分離性能を向上させることが報告されている。
- ◆本研究では、学習データから ILRMA 内の基底行列を事前に作成する。基底ベクトルは単音のスペクトログラムの低ランク近似により作成される。更にインパルス応答のスパース性も考慮した手法を提案する。
- ◆客観指標を用いた数値実験によって提案法の優れた分離性能を示す。

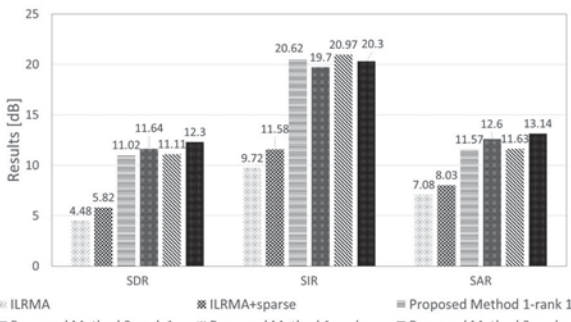


Fig. 1: Averaged separation performance for mixed sounds of piano and bass

### 1-Q-26

#### 1-Q-26 血液透析患者のシャント雑音解析に基づくバスキュラーアクセス機能評価法と超音波エコー法との比較検討

Comparison and examination between ultrasonic pulse-echo technique and vascular access function evaluation method based on shunt murmur analysis of hemodialysis patients

○佐々木一真, 関 俊輔, 佐藤敏夫(桐蔭横浜大院・工学研)

奥 知子, 本橋由香, 山内 忍(桐蔭横浜大・医用工)

血液透析患者のバスキュラーアクセス (VA) を模擬した擬似血管狭窄モデルを作製し、電子聴診器で擬似シャント音を採取した。その擬似シャント音に対してウェーブレット変換による時間-周波数解析を行い、正規化相互相関係数 R と正規化持続時間 NDT を算出した。この R と NDT による VA 機能評価法と、超音波エコー法から得られる血管抵抗指数 (RI) による VA 機能評価法との比較検討を行った。

電子聴診器で採取した擬似シャント音から得られる R と NDT は、超音波エコー法で得られる RI との間に強い相関関係があることがわかった。血液透析患者から得られた結果からも、シャント音から算出した R や NDT と、超音波エコー法で得られた RI との間には相関関係があることが確認できた。

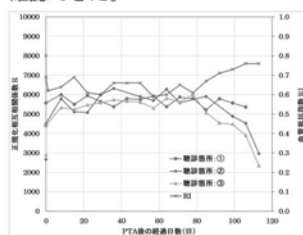


Fig.1 Change with time of R and RI of hemodialysis patient

### 1-Q-25

#### 1-Q-25 嚥下音の時間-周波数解析に基づく嚥下リハビリテーション効果の定量的評価の試み

Quantitative evaluation of swallowing rehabilitation effect based on time frequency analysis of swallowing sound

○可児雅弥, 佐藤敏夫(桐蔭横浜大院・工学研)

奥 知子, 本橋由香, 山内 忍(桐蔭横浜大・医用工)

我々は、頸部から採取した嚥下音に対してウェーブレット変換による時間-周波数解析を実施してきた。また、嚥下音信号の包絡線を3次スプライン近似で作成し、3つの音の時間的位置関係を表すパラメータ SFP (Swallowing Function Parameter) を算出することで、嚥下機能を評価することを検討してきた。今回の報告では、嚥下反射の惹起性を高めるとして注目されている冷却炭酸水および嚥下調整食として用いられるゼリー食について、各試料嚥下時に得られた SFP 値に基づいて、その有効性や最適性について検証した。まず、冷却炭酸水については、炭酸強度が大きくなるほど SFP 値が低下し、冷却炭酸水は嚥下反射の惹起性向上に有用であると考えられた。また、ゼリー食やとろみ食などの嚥下調整食については、硬さが約 4150 N/m<sup>2</sup> の水ゼリー嚥下時に SFP 値に特徴的な低下が見られた。

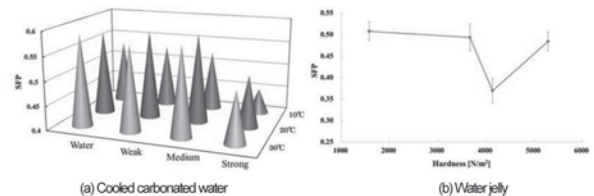


Fig.1 SFP calculation results for cooled carbonated water and water jelly

### 1-Q-27

#### 1-Q-27 上腕上部下部同時 MMG 計測による筋活動計測の一検討

The clamped brachium MMG measurement by the upper/lower measurement point for muscle activity monitoring

☆大里一矢, 山下涼介(NBU), △石川直生, 上原正志(CIT),

△樋口幸(大分県立看護科学大), △市田秀樹(信州大学),

福島学(NBU), △森竹隆広(シェルエレクトロニクス(株)),

風間道子(NH ラボ), 柳川博文(arsl)

- ◆MMG を用いた筋活動把握の検討

上腕筋上部(ACC1)下部(ACC2)同時計測

・伸縮活動: 計5動作を対象

動作: 縮, 維持, 伸, 着机前, 着机

・対象筋繊維群: 計3群

筋繊維群: 上腕二頭筋, 上腕三頭筋, 上腕筋



Fig.1 対象筋繊維群の構図図

[森竹隆広, 解剖学第1巻: 新説・骨学・理学療法・筋学, 改訂第11版, pp.346, 2002]

- ◆対象筋から離れたセンサ位置での計測可能性を確認

- ・伸縮活動: ①縮(▽), ②維持(○), ③伸(△), ④着机前(□), ⑤着机(◇)
- ・筋繊維群: 上腕二頭筋(x), 上腕三頭筋(+), 上腕筋(・)

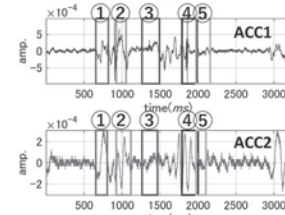


Fig.2 MMG 時間波形

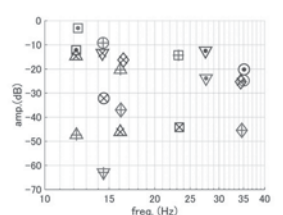


Fig.3 筋伸縮動作と筋繊維群の対応関係

### 1-Q-28

#### 1-Q-28 ドローン飛行音を音源とした音響測距法の可能性調査

A case report of an reflector detection with the acoustic distance measurement method using drone flight sound

☆山下涼介, 大里一矢(日本文理大学), 坂口智弘, 上原正志(千葉工大), 福島学(日本文理大学), △森竹隆広(シェルエレクトロニクス(株)), △小野俊二(シーアイロボティクス(株)), 風間道子(NH ラボ), 柳川博文(arsl)

◆音響測距法【鈴木, 他, IEICE, Vol. J97-A, No.4, pp.343-pp.346, 2014】

観測信号のワースペクトルに生じるコサイン変調から距離を推定  
$$p(f) \approx A^2 \left\{ 1 + 2 \sum_{n=1}^M \gamma_n \cos \left( \frac{2\pi f d_n}{c} + \phi_n \right) \right\}$$
  
但し,  $f$ は周波数,  $\gamma_n$ は反射係数,  $\phi_n$ は反射位相,  $A$ は音源信号の振幅,  $d_n$ はマイクと反射物との往復距離,  $c$ は音速

◆ドローン飛行音が音源として使えるか?

モデル検証実験 計測に成功

実機検証実験 半径 30 (m) (航空法)

計測に成功

物理位置と推定値が対応

Fig.1 実験環境俯瞰図

Fig.2 推定値と地形断面図

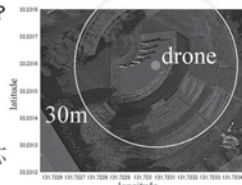


Fig.1 Overview of the measurement place

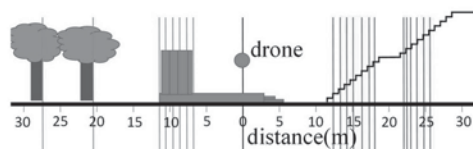


Fig.2 Correspondence between estimated distance and topographical cross section

### 1-Q-30

#### 2 台のパラメトリックスピーカの配置による音像定位の検討 -上方配置について-

Study on sound localization by arrangement of two parametric loudspeakers -Arrangement in upper region-

☆福田隆人, △古田昇輝, 青木茂明, 島内末廣(金沢工大)

◆以前の検討で最も優れた定位感を示したスピーカ配置(方位角±60°, スピーカと聴取者の距離 120 cm)において, スピーカの仰俯角と前方方向の音像定位との関係を検討した。

◆試験には HRTF を畳み込んだピンクノイズを使用し, スピーカは聴取者から見て方位角 60° の方向に 120 cm 離し, 仰俯角 +30°, 0°, -10°, -30° となるように配置を変更した。

◆仰俯角 +30° 配置と -30° 配置のときの定位方向の分布図を Fig. 1 に示す。音像は全体的に HRTF の角度の方向に定位した。正答率の差は約 10 % であり, 仰俯角を +30° ~ -30° の範囲で変更しても前方の音像定位に大きな変化はないことがわかる。しかし, 個人差が見られる角度もあるため最適なスピーカ配置が個人によって違う可能性も考えられる。

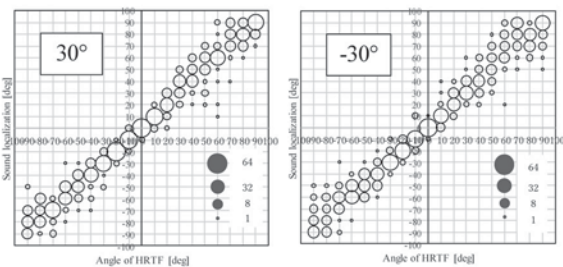


Fig.1 Localization distribution

### 1-Q-29

#### 1-Q-29 CNN を用いた船舶航行音による種別分類方法の検討

A study on boat noise recognition for boat type categorization using convolutional neural network

☆山口晴己, 武藤憲司(芝浦工大)

◆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。

◆船舶の検出においては 5 秒間の音分析を行ったとき, 十分に船の検知を行うことができた。しかし船種分類においては十分な分析時間長ではなかった。

◆本報告では航行音による船種分類方法を検討するため, CNN モデルにスペクトログラムを利用した学習を行い, 船種種別に必要な学習要素を検討した。本報告では学習要素のうち船の種別が人でも十分に行えた 60 秒の分析時間における結果を報告する。

◆CNN モデルの学習より 1 時間の環境音に対し船種分類を行った結果を Fig.1 に示す。10 隻中 7 隻が所望する分類と同一の結果となった。

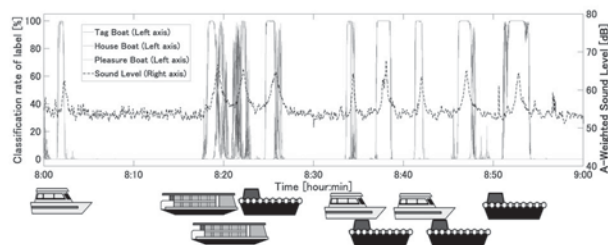


Fig.1: Result of boat type classification by CNN model.

### 1-Q-31

#### 1-Q-31 RBF カーネル SVM を用いたシャント音血管狭窄検出における特徴量算出法

A method of feature quantity calculation for shunt murmur angiostenosis detection using RBF kernel SVM.

☆東大輔, △野田史也(大分大), △田中佳子, △秦さと子(大分科大), 西島恵介, 古家賢一(大分大)

◆腎不全などの患者は日々のシャント管理が求められるが, シャント機能の判断は主観に依存してしまうため, シャント機能を手軽に自動で判断するシステムが求められている。

◆本報告では, 従来手法で行われたシャント音の固定長切り出しではなく, 一脈拍に対応する可変長の切り出しを行う。また, 一脈拍の切り出しによって患者ごとにシャント音のピークにずれが生じるため, 特徴量算出法の改善を行う。さらに, 識別器として RBF カーネルを用いた SVM の検討を行う。

◆実験結果より, 提案手法を用いることで精度の向上が確認された。また, カーネルの変更によって精度の向上が確認された。

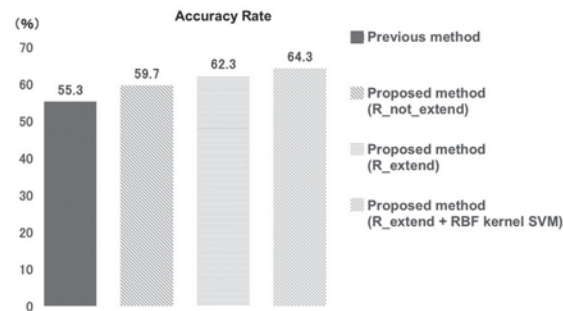


Fig.1: The result of SVM detection

### 1-Q-32

#### 1-Q-32 改良型ランダム素子配置 パラメトリックスピーカの開発

Development of improvement parametric loudspeaker with random transducer arrangement.

☆宮澤佑樹, 金城翔太, 藤岡豊太, 永田仁史(岩手大)

- ◆ 指向性のメインローブの近くに出現するサイドローブを抑え、さらに密な素子配置による出力パワーの向上も考慮し配置する素子の分布半径を小さくしたランダム素子配置パラメトリックスピーカを開発した。
- ◆ 超音波振動子の出力特性のばらつきを考慮し素子1つつつ音波を出力、録音し各素子の応答特性が近いものを選んだ。
- ◆ 各ステアリングの指向性を測定し既開発のパラメトリックスピーカと比較検討し、同時に出力パワーの測定も同時に行なった。
- ◆ ステアリング0°の時既開発スピーカで出現していた副極が新規開発スピーカでは抑えられているわかった。
- ◆ 再生音の出力に関しては既開発スピーカに比べて3.6 dB程出力音圧レベルが向上した。

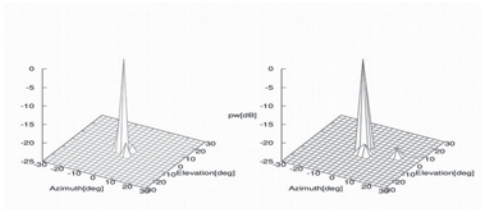


Fig.1: Directivity of demodulated signal at azimuth 0°(random transducer arrangement speaker)

### 1-Q-33

#### 1-Q-33 ステレオカメラを用いたパラメトリックスピーカの 制御インターフェースの検討

A study on control interface of parametric loudspeaker using stereo camera

☆藤原真也, 金城翔太, 藤岡豊太, 永田仁史(岩手大)

- ◆ ステレオカメラを用いてパラメトリックスピーカの照射位置を指示するシステムを開発した。
- ◆ ステレオカメラの性能を評価するためにステレオカメラから1.5mと2.7mに仮想照射点を設置し、三次元座標をカメラで測定し真の位置と計測値の誤差を比較した。
- ◆ Fig1に奥行き1.5m地点に関する真の位置とステレオカメラによる計測結果、Fig2に奥行き2.7m地点に関する真の位置とステレオカメラによる計測結果を示す。
- ◆ 奥行き1.5mと2.7mに関しての真の位置とカメラ画像から計算した位置との誤差は小さく、パラメトリックスピーカの照射点指示を補助するのに十分な性能であるといえる。

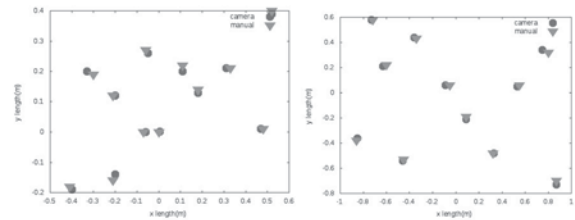


Fig.1 Estimation result (depth 1.5m) Fig.2: Estimation result(depth 2.7m)

### 1-Q-34

#### 1-Q-34 音響信号を用いた 空間センシングに関する予備的検討

A preliminary study of object detection system using acoustic signal.

◎金城翔太, 藤岡豊太, 永田仁史(岩手大)

- ◆ パラメトリックスピーカを用いた空間センシングについて検討した。
- ◆ 反射のある環境において、ステアリングビーム方式と合成開口方式、M系列信号を用いる従来手法の3手法で実験を行った。
- ◆ ステアリングビーム方式と合成開口方式において、物体の検出ができた。

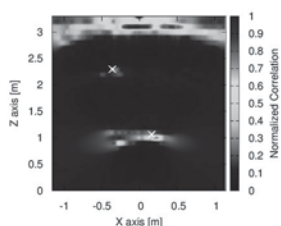


Fig.1: Result of steering beam method.

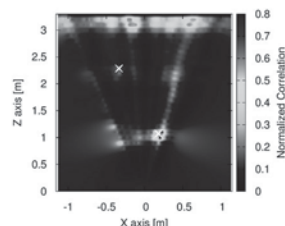


Fig.2: Result of synthetic aperture method.

### 1-Q-35

#### 1-Q-35 等価点音源を用いた局所領域への 直接音伝達関数の推定

Estimation of direct-sound transfer functions to local area using equivalent sources

☆津國和泉, 田中琴美, 黒川翔瑠, 池田雄介, 小坂直敏(東京電機大)

- ◆ 問題点
  - 効率的な多点のインパルス応答計測は高コスト
- ◆ 提案手法
  - 目的: 少数点の計測よりスピーカからの直接音伝達関数を推定
  - スピーカを取り囲む閉曲面上に等価音源(多重極音源)を配置
  - スペース最適化によって等価音源の重み係数を決定
- ◆ 自由空間における三次元音場シミュレーション実験
  - 5chの線形マイクロホンアレイを使用
  - モノポール音源+ダイポール音源によって、より広範囲の推定が可能
  - 評価領域 (Fig.1 白線内部) では -40 ~ -50 dB 程度の誤差

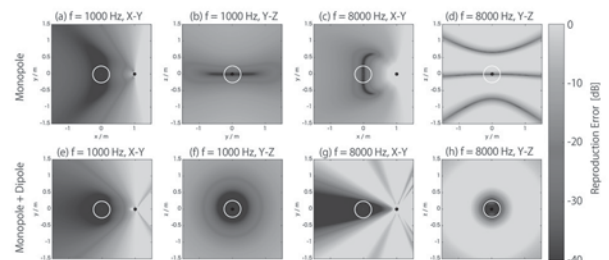


Fig.1: Error of the estimated transfer function using monopole and dipole sources (Evaluation area are inside a white circle.)

### 1-Q-36

#### 1-Q-36 等価点音源による局所音場への一次反射音伝達関数の推定

Estimation method of transfer functions of primary reflection for local sound field by equivalent point sources.

☆田中琴美, 津國和泉, 池田雄介, 小坂直敏(東京電機大)

##### ◆問題点

初期反射を考慮した伝達関数の推定には複雑なマイクロホンアレイ及び多数のマイクロホンが必要

##### ◆提案手法

- 目的: 少数計測点から一次反射音を含む伝達関数のモデル化
- ・スピーカとその鏡像の周囲に仮想的な等価点音源を複数配置
- ・スパース最適化により最適な点音源の重み係数を決定

##### ◆無響室にて評価実験

- ・一面の音響反射板を用いて単一の一次反射音を含む伝達関数を推定
- ・全周波数帯域で-35.4 ~ -3.4 dB 程度の誤差

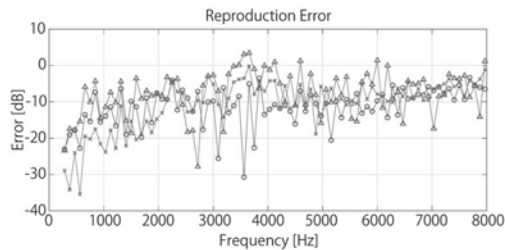


Fig.1: Reproduction Error in proposed method

(○: -0.2 m on the X axis, ×: -0.1 m on the Y axis, △: -0.2 m on the Y axis)

### 1-Q-38

#### 1-Q-38 ライン露光順次読み出し方式カメラを用いた撮影動画像に基づく音声獲得

Speech Capture Based on Movies with Rolling-shuttered Visual Camera

☆寺野光一, 進藤広暉(立命館大院), 福森隆寛, 西浦敬信(立命館大)

音は疎密波として空气中を伝搬するため、音源付近に存在する物体の表面は音波によって振動している。我々はこれまで、ライン露光順次読み出し方式カメラを用いて物体の表面振動を撮影した画像から音波を抽出する手法を提案してきた。本稿では、この手法を拡張し、より長時間の音源収録を可能にするため、ライン露光順次読み出し方式カメラを用いて撮影した動画像から音声を獲得する手法を提案する。提案法は、撮影した動画像に対して、1フレーム毎に音波抽出手法を適用し、抽出した信号を時系列順に接続することで音声を獲得する。また、提案法の有効性を確認するため、スピーカの振動子に貼付したアルミシート上の直線を振動物体として音声獲得実験を実施した。実験の結果、Fig.1に示す通り、提案法によりライン露光順次読み出し方式カメラを用いた撮影動画像から1 kHz以下の周波数帯域において音声を獲得できることを確認した。

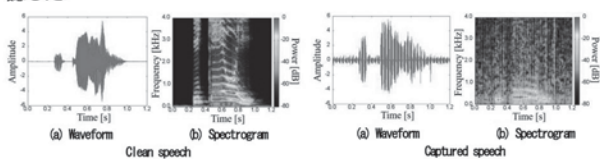


Fig.1: Experimental result.

### 1-Q-37

#### 1-Q-37 手持ちマイクロホン用いた動的インパルス応答における計測位置誤差の影響

Effect of measurement position errors on dynamic impulse response measurements using hand-held microphones

☆渡辺靖明, 池田雄介(東京電機大),

片岡優太, 及川靖広(早大理工), 小坂直敏(東京電機大)

##### ◆これまで手持ちマイクロホンと複合現実技術を用いた3次元音響インテンシティ可視化システムを提案

高度化への問題点: 測定位置誤差の存在

##### ◆動的インパルス応答計測への適用を目的とし、システムの計測位置誤差を評価し、誤差がインパルス応答の復元に与える影響を調べた。

##### ◆実験により、極端に誤差が大きくなる場合を除いて、約0.05 m程度の測定位置誤差があることがわかった (Fig. 1)。

##### ◆シミュレーションにより、0.05 mの誤差がインパルス応答の復元に影響を与えることが示唆された。

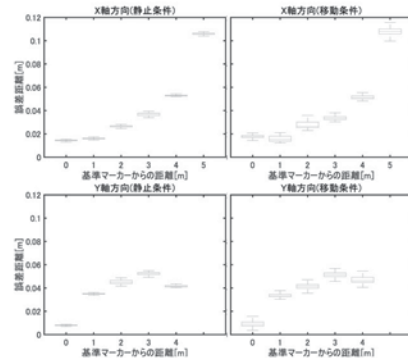


Fig.1: Measurement position error

### 1-Q-39

#### 1-Q-39 デジタルスチルカメラによる撮影画像を用いた抽出音波の帯域拡張

Band extension for captured sound using an image with digital still camera

☆進藤広暉, 寺野光一(立命館大院), 福森隆寛, 西浦敬信(立命館大)

近年、ハイスピードカメラ撮影による高品質動画を用いた音波抽出手法が提案されている。この手法は、高フレームレート動画におけるフレーム間の画素の変位から物体の振動を抽出し、物体に到来した音波を得る。しかし、高フレームレートの映像が得られない状況では可聴域の音波を抽出することは困難であった。この問題に対し、我々はComplementary metal-oxide semiconductor (CMOS) センサ搭載のデジタルスチルカメラで撮影した低品質動画像から音波を抽出する手法を検討している。しかし、本手法は被撮影物体の振動が微小となる高周波数成分の抽出が困難であり、従来法では0.8 kHzのまでの周波数成分の抽出が限界であった。そこで本稿では、撮影環境及び画像のエッジ検出手法を検討することにより抽出音波の帯域拡張を目指す。Fig.1に1.0 kHzの純音を抽出した結果を示す。(a)に撮影画像、(b)にエッジ検出における二値化画像、(c)に抽出音波の波形、(d)に抽出音波のパワースペクトルをそれぞれ示す。Fig.1から、提案法により従来法よりも高周波数の音波が抽出できていることを確認した。

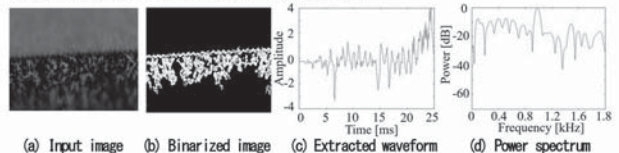


Fig.1: Experimental result.

### 1-Q-40

#### 1-Q-40 空間標本化定理の拡張に関する考察

Study on extending the spatial sampling theorem

☆井上和弥, △鈴木遼平, △大内翔, 島内末廣, 青木茂明(金沢工大)

- ◆空間標本化定理によれば、マイクロホンアレイの素子間隔を波長の1/2以下にしなければ、正しく音場の情報を獲得することができない。
- ◆本報では、音圧のみではなく導関数も標本化することによりマイクロホンアレイの実効的素子間隔の拡大を試みる。
- ◆シミュレーション実験により、音圧のみの標本化(Fig.1)と比べ、音圧と導関数の標本化(Fig.2)により、空間エリアシングを回避可能であることを確認した。

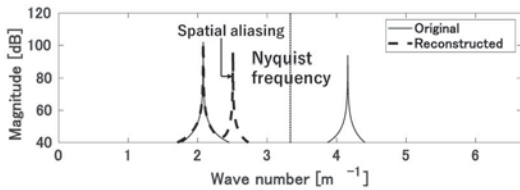


Fig.1 Reconstruction result based only on sound pressure samples.

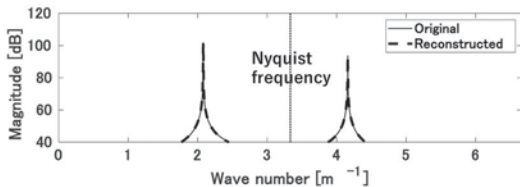


Fig.2 Reconstruction result based on sound pressure and its first derivative samples.

### 1-Q-42

#### 1-Q-42 ミスナー法を用いた格子点音圧データからの球面調和係数の推定

Estimation of spherical harmonic coefficients from sound pressures on the grid points based on Misner's method

☆松岡 良憲, 大谷 真(京都大学大学院工学研究科)

- ◆Misnerは、重力波の測定・分析の分野において、格子点上データからある球面上での球面調和係数を推定する手法(以下、ミスナー法)を提案した。
- ◆本稿では、このミスナー法を音圧場に適用することにより、格子点音圧データから球面調和係数を推定し、高次アンビソニックスなどの音場再現理論に応用する手法について数値的検討を行った。
- ◆本稿の解析条件では、0~2次において高い精度が保証されており、3次、4次においても格子点間隔を適切に定めることで高い精度で推定できることがわかった。
- ◆格子点や球面の配置が既知であれば計算は大幅に省略され、格子点データの値に依らず短時間で推定が可能となることも確認された。

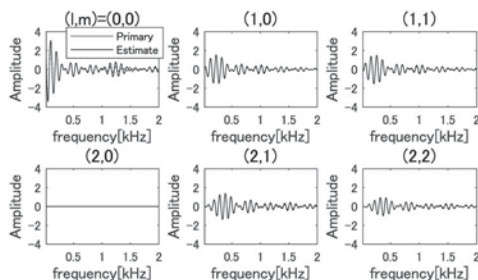


Fig.1: spherical harmonic coefficients of order 0~2, grid interval h=0.1 [m] (real part, frequency domain)

### 1-Q-41

#### 1-Q-41 高解像度録音を用いたベルカント唱法による歌声の広帯域音響特性解析

Broadband acoustic analysis of Bel canto singing voice using high resolution recording

○中里尚子, 大久保寛(首都大院・システムデザイン)

- ◆様々な歌手の歌声を計測することで、プロの歌手など習熟度の高い被験者とその他のアマチュア・初心者被験者との歌声の音響特性の違いについて検討を行った。
- ◆また、計測に 32bit/768kHz 高解像度録音システムを用いているため 100kHz までの解析が可能であることから、歌声の中の可聴域外の周波数成分の多寡について考察した。
- ◆その結果、プロ被験者はアマチュア・初心者被験者と比べ歌声に多くの倍音があり、また高音域の歌声似は可聴域外の周波数帯域の成分も含まれている事が確認できた。

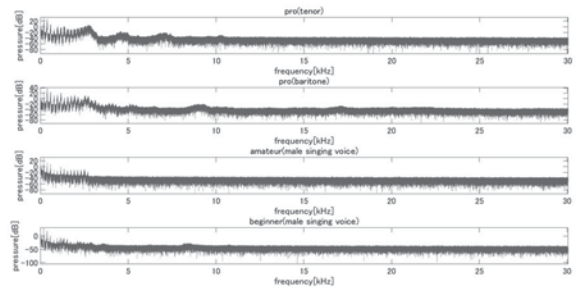


Fig.1 Frequency spectrum of male humming voice when he hums with "mi" of G3.

### 1-Q-43

#### 1-Q-43 スピーカ・マイク一体型球面アレイを用いた反射位置探索

Reflected position search using spherical array with integrated loudspeakers and microphones.

◎小林 真萌, 羽田 陽一(電通大)

- ◆スピーカ・マイク一体型球面アレイを用いた室内の反射位置探索手法を提案する。
- ◆Fig. 1 にシステム構成を示す。測定したインパルス応答から方向別インパルス応答に変換し、ビームの角度分解能を考慮しながら反射音成分を抽出した。
- ◆実際の部屋での実験では、壁の位置と特定した反射位置が概ね一致していることを確認した。

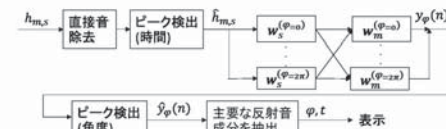


Fig. 1: System block diagram.

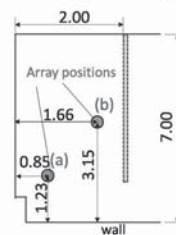


Fig.2: Experimental condition.

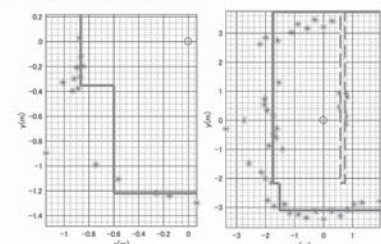


Fig.3: Estimated reflection positions.

### 1-Q-44

#### 1-Q-44 動的モード分解を用いた自動車内データの変化検知

Change detection using dynamic mode decomposition for time series sensor data in vehicle.

○伊藤弘章, △村田伸, 原田登(NTT)

- ◆自動車の車両損傷イベント検知のための、自動車内データを用いた観測データ駆動型の変化検知手法を提案する。
- ◆従来の特異スペクトル変換を用いた変化検知手法では、静的な特徴パターンにより変化度を算出するため、運転時の時間変化を捉えきれず、検知精度が低い。
- ◆本手法では、動的モード分解を用いて観測データの時間発展に由来する特徴を抽出することで、変化度を算出する。
- ◆実環境で収録した車両損傷イベントデータを用いた変化検知実験により検知性能を比較し、本手法の有効性を確認した。

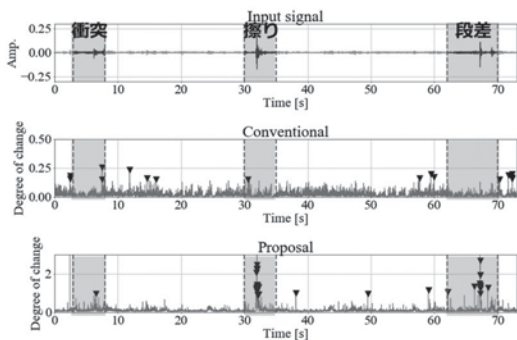


Fig.1: 変化度算出結果の比較。

灰色網掛け部はイベント区間、▼は閾値処理による変化検出点を示す

### 1-Q-46

#### 1-Q-46 SDD 法による非線形2次元アレイスピーカの駆動

Driving of nonlinear two-dimensional array speaker by SDD method

☆田代康平, 田村安孝, 柳田裕隆(山形大院)

- ◆非線形の特徴を持つ電気音響変換素子を用いたエッジ接続型2次元アレイは、少ない信号チャンネルで駆動できる。これまで提案された駆動方式は、動作周波数や量子化誤差が大きいなどの問題があった。
- ◆量子化誤差の低減を目的に、駆動信号を行列分解の手法であるSDDと $\Delta\Sigma$ 変調により発生させる方法を検討した。
- ◆シミュレーションにより、 $\Delta\Sigma$ 変調を適用した場合としない場合の比較、PWM駆動するための動作周波数の決定を行った。
- ◆シミュレーションの結果、駆動信号を発生できることが確認できた。また、 $\Delta\Sigma$ 変調を適用した場合は高調波のノイズが低く抑えられていることがわかる。PWM駆動するための動作周波数は、約800kHzとなった。

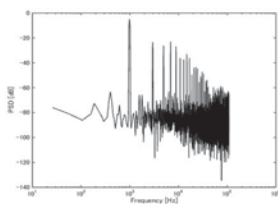


Fig.1: Power spectrum when ΔΣ modulation is not applied

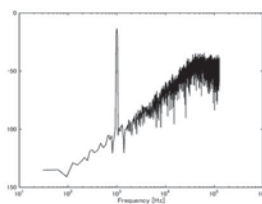


Fig.2: Power spectrum when ΔΣ modulation is applied

### 1-Q-45

#### 1-Q-45 充電電池の音色で残量の有無を見極める方法

Effective Method Using Acoustic Characteristics for Screening Discharged Rechargeable Battery

☆馬込智瑛, 大久保寛(首都大院・シス)

- ◆乾電池・充電電池の残量の有無を外見や重量から判断することは難しい
- ◆乾電池: それらの打音を聴くことで残量の有無を判別できる
- ◆充電電池: それらの打音を聴くだけで判別することは難しい
- ◆機械学習 (SVM) による判別が可能かどうかを検証した

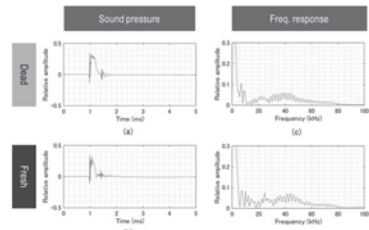


Fig. 1: Examples of SP signal and freq. response of hammering sound of dead / fresh rechargeable batteries

Output Class	Target Class Empty		Target Class Full	
	Empty	Full	Empty	Full
Empty	493 49.3%	81 8.1%	10 1.0%	98.0%
Full	7 0.7%	419 41.9%	490 49.0%	3.9%
	98.6% 1.4%	83.8% 16.2%	91.2% 8.8%	97.0% 3.0%

$\Delta F = 0.1 \text{ kHz}, D = 200$

Output Class	Target Class Empty		Target Class Full	
	Empty	Full	Empty	Full
Empty	480 48.0%	10 1.0%	20 2.0%	98.0%
Full	20 2.0%	490 49.0%	490 49.0%	3.0%
	96.0% 4.0%	98.0% 2.0%	97.0% 3.0%	97.0% 3.0%

$\Delta F = 0.1 \text{ kHz}, D = 1000$

Fig. 3: Confusion matrix of screening the residual quantity of rechargeable batteries by trained classifier; (a) 0 Hz - 20 kHz; (b) 0 Hz - 100 kHz

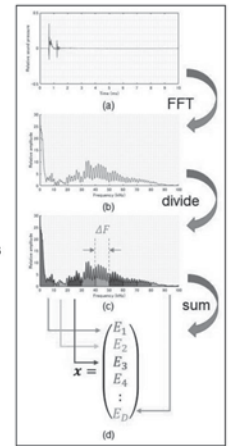


Fig. 2: How to generate feature vectors to train classifier for screen dead rechargeable batteries

### 1-Q-47

#### 1-Q-47 PVDF 圧電体 d33 の真の値

PVDF 圧電体の d33 の値とポアソン比

○森山信宏(クレハ), 須藤隆一, 小川智幸(エルメック電子工業)

- ◆PVDF 圧電体の厚さ方向の圧電ひずみ定数 d33 の測定方法において、前報(2018 年秋)にて、押圧具とサンプルの間にサンプルと同等の弾性挙動を持つ相応の厚さの材料を挿入する事により、横方向の束縛力の影響を無くして、真の値を測定できることを提案した。(Fig. 1, Fig.2)

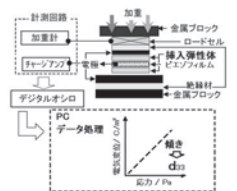


Fig.1 The schematic diagram of d33 measurement

- ◆この事から、

$$d_{33} / d_{33}^M = 1.818$$

が得られ、束縛力の関係

$$d_{33}^M = d_{33} + (d_{31} + d_{32}) \cdot \nu_p / (1 - \nu_p)$$

から

$$\nu_p = (d_{33}^M - d_{33}) / (d_{31} + d_{32} + d_{33}^M - d_{33})$$

が得られる。

d31, d32 定数を測定し、

ポアソン比を求めた。

Table 1 測定データ及びポアソン比

Sampling position	A	B	C	Average
d31 pC/N	28.1	27.9	28.6	28.2
d32 pC/N	0.2	0.5	1.0	0.6
d33M pC/N	-18.7	-17.1	-18.8	-18.4
*d31 = 1.818d31 ⇒ pC/N	-34.0	-32.1	-34.1	-33.4
Poisson's ratio νp	0.3502	0.3375	0.3414	0.3430

注\*: d31/d33 = 1.818 森山, 須藤 音響論 (報), 321-322, 2018<sup>1)</sup>

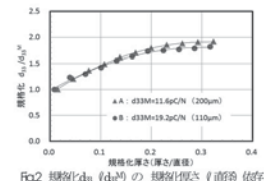


Fig.2 規格化d33 (d33^M) の規格化厚さ(直筋)依存性

### 1-Q-48

#### 1-Q-48 大気圧コロナ放電音のインパルス応答

Impulse response of corona discharge sound in atmospheric pressure  
 ☆岡崎凌大, 奥山由, 工藤彰洋(苫小牧工業高等専門学校)

◆大気圧コロナ放電を利用した Plasma loudspeaker は、Dynamic type loudspeaker と異なり振動板を持たず、幅広い周波数特性をもつ。そこで、Plasma loudspeaker と Dynamic type loudspeaker のインパルス応答を測定した。Fig.1 に測定したインパルス応答をフーリエ変換した結果を示す。同図より 60 kHz 未満では(a)が $-3.4$  dB/oct., (b)は $-4.0$  dB/oct., 60 kHz 以上では(a)が $-4.0$  dB/oct., (b)は $-11.3$  dB/oct.となり、Plasma loudspeaker の方が Dynamic type loudspeaker よりフラットな特性であった。

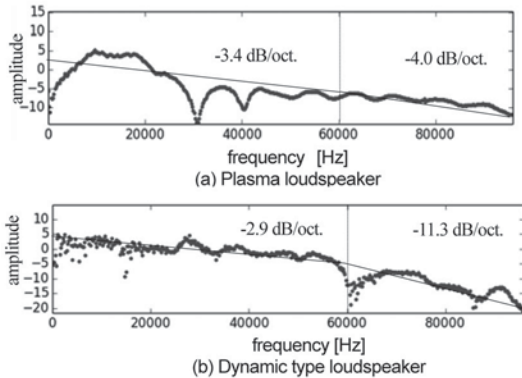


Fig.1: Fourier transform of the impulse responses

### 1-Q-50

#### 1-Q-50 時間反転波を用いた局所的空中音響通信の数値シミュレーション

Numerical Simulation of Local Acoustic Communication Using Time-reversal Wave  
 ☆本井隆輝(筑波大院・シス情工), 海老原格, 水谷孝一, 若槻尚斗(筑波大・シス情系)

◆時間反転波を用いることで、受信機の位置に音波を集束させることで、ユーザに不快感を与えない局所的音響通信の実現可能性を検討した (Fig. 1)  
 ◆室内通信を仮定したシミュレーションの結果、提案手法は時間反転波を利用しない既存手法と比較して、音圧分布が局在し (Fig. 2) 通信品質も向上していた。

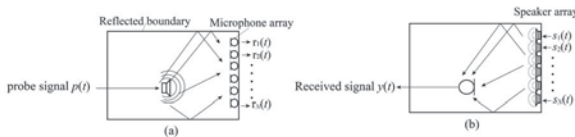


Fig. 1 Simulation environment

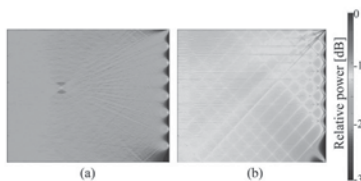


Fig. 2 Sound pressure distribution obtained from simulation; (a) proposed method and (b) conventional method

### 1-Q-49

#### 1-Q-49 スマートフォン向け小型音響機器の自動設計支援システムに関する一検討

A Study on Automatic Design Support System for Compact Acoustic Devices  
 ☆奥西洋貴, 平井開, 梶川嘉延(関西大学)

◆スマートフォンに搭載される小型音響機器は、近年の小型化・複雑化に伴い従来の解析法であった音響等価回路解析法が適用できず、熟練の技術者によるカットアンドトライ設計に頼るしかないという問題点が存在する。この問題を解決するために、我々はディープラーニング(DNN)を用いた自動設計支援システムを提案し、検討を行ってきた。  
 ◆提案システムでは DNN を用いることで筐体の寸法値と対応する周波数特性との関係を自動的に学習する。学習を行ったシステムに所望の周波数特性を入力することで、所望の特性に対応する寸法値を得ることが可能であった。  
 ◆昨今のスマートフォンに搭載される小型音響機器の構造は筐体構造に対する制約のため、これまで検討してきた音響構造と異なり振動板が音の伝搬方向に対して横向きに設置されている。本稿ではこの異なる構造に対しても提案した自動設計支援システムが高い精度で設計可能かどうかについて検討を行う。

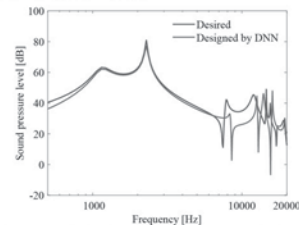


Fig. 1: Comparison between desired frequency response and designed one by DNN

### 1-R-1

#### 1-R-1 Frequency characteristics of threshold of cartilage-conduction by using a cellphone panel with a PZT-based vibrator

© Xiuyuan Qin (Chiba Univ.), Sho Otsuka (Chiba Univ.), Junghun Noh (Samsung Display), and Seiji Nakagawa (Chiba Univ.)

◆Cartilage-conduction (CC) has been proposed and applied to a hearing aid for atresia of the external auditory canal. Recently applications of CC on mobile phone have been developed.  
 ◆In this study, hearing thresholds of CC were measured with different given placements by using a cellphone panel. A comparison between CC and air conduction (AC) was made, and the influence of placement on CC was discussed.  
 ◆The measurements were done when the panel contacted or did not contact the pinna as well as the panel covered the pinna fully and partially.  
 ◆The hearing thresholds on the contacted condition were lower than the results obtained on the uncontacted condition, showing the influence of CC by vibration of mobile phone.



Fig. 1: Fully Covered

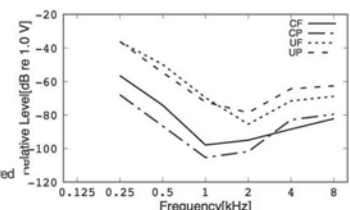


Fig. 2: Partly Covered

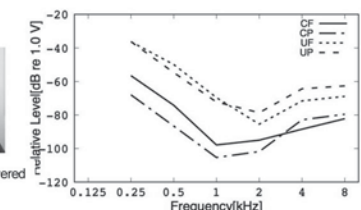


Fig. 3: The hearing thresholds measured on four different conditions.

### 1-R-2

#### 1-R-2 通常発声とささやき声を対比した寸法知覚の計算モデル

A computational model of speaker size perception for when comparing voiced and whispered speech sounds

☆上村 怜央, 入野 俊夫(和歌山大), ロイ D. パターソン(ケンブリッジ大)

- ◆有声音(Vo)とささやき声(Wh)を音声刺激対とした寸法弁別実験 VTLの異なる基準音と比較音を対(Vo-Vo, Vo-Wh, Wh-Wh, Wh-Vo)で聞かせ、どちらの音声が「小さい寸法」の話者に聞こえるか回答。
- ◆被験者実験の結果と計算モデルによるシミュレーション結果 Fig.1に、基準音がささやき声(Wh)でVTLが小さい領域の結果を示す。Fig.1(a)の被験者実験において、Wh-Wh(点線)とWh-Vo(実線)の心理物理曲線の形状は大きく異なっている。Fig.1(c)のシミュレーション結果では、Wh-Vo(実線)の心理物理曲線の中心は大きくシフトしており、再現性は高くないようにみえる。一方、Fig.1(b)で、重み付け関数 SSI weight を導入すると、Wh-Wh(点線)でも Wh-Vo(実線)でもほとんど同じ形状になった。しかし、Wh-Vo(実線)は被験者実験の結果に比べて明らかに傾きが急峻であり、さらなる検討が必要である。

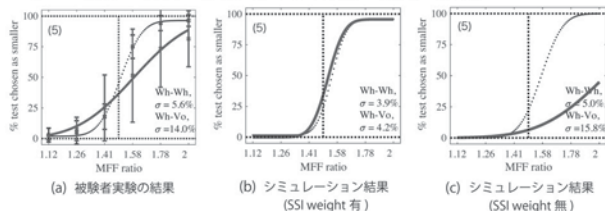


Fig. 1. (a) Psychometric functions for size discrimination experiments by human listeners (b) Simulations of psychometric functions from the model when it includes "SSI weight" (c) Simulations of psychometric functions from the model when it does not include "SSI weight"  $\sigma$ : just noticeable difference(JND) in percent. Dotted line: a pair of whispered speech. Solid line: combination of whispered and voiced speech.

### 1-R-4

#### 1-R-4 両耳性の遅延パルス列加算によるオクターブ・シフトに対する両耳情報の影響

The octave-shift impression by summation of two desynchronized pulse trains and its relation to binaural mismatch.

☆渡邊 菜, 津崎 実(京都市芸術大学)

- ◆同一の周期を持つ二つのパルス列を周期の1/2に相当する遅延を加えて表示すると原周期よりも1オクターブ高く知覚される(オクターブ・シフト)。
- ◆同一の周期を持つ二つのパルス列が異なる方位角にあると仮定してそれぞれに両時間差を加え、さらに2音間に開始時間差を与え、結果として片耳にはちょうど半周期の、反対耳には半周期からずれたパルス列が提示され、**両耳に提示される周期が矛盾する状況を作った。このときオクターブ・シフトは生じるのか?**原周期のパルス列のピッチと半周期パルス列のピッチそれぞれがどの位よく聞こえるかを評定尺度法で判断。
- ◆片耳に半周期、反対側の耳に半周期からずれたパルス列を両耳提示すると、半周期に相当するピッチも原周期に相当するピッチも単耳提示の場合よりも聞こえにくくなるものの知覚され、どちらか一方のピッチに集約されるという結果にはならなかった。

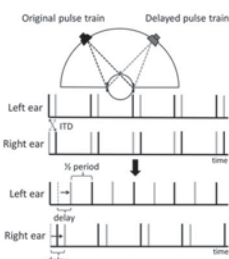


Figure 1 The conceptual diagram of the test stimulus.

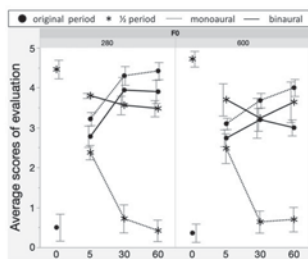


Figure 2 The average scores of the evaluations.

### 1-R-3

#### 1-R-3 音声帯域における骨導音の順向性マスキング効果の周波数依存性

Frequency dependency of forward masking effects in voice band using bone conduction headphones

○碓井 峻, 岩城 護(新潟大院・自然研)

- ◆骨導音聴取時に生じる順向性マスキング効果は1000, 2000, 4000 Hz 純音に対して調査されてきた。本稿では、500 Hz も含めて、音声帯域における周波数依存性を調査した。
- ◆ノッチ雑音マスキング法を用いて、骨導音聴取時と気導音聴取時の純音に対するマスキング量を計測した。純音の周波数は500, 1000, 2000, 4000 Hzの4種類とした。
- ◆骨導音聴取時でも順向性マスキング効果が確認された。また、気導音聴取時とは異なる特性が確認された。
- ◆骨導音聴取時のマスキング量は、1000, 2000 Hzに比べて、500, 4000 Hzでは多くなる傾向が見られた。特に4000 Hzでは顕著だった。

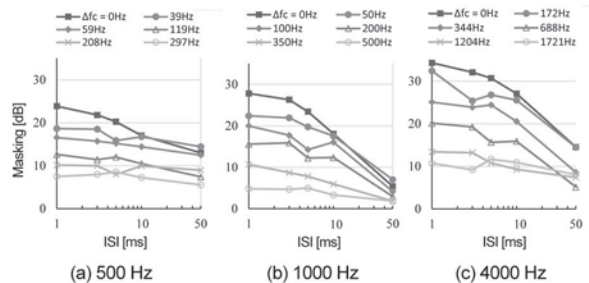


Fig. 1: Forward masking characteristics for bone conduction hearing.

### 1-R-5

#### 1-R-5 2純音の同時性知覚における音圧と周波数領域の影響

◎岡崎 聡, 津崎 実(京都市立芸大)

- ◆2音が同時に開始したと知覚される同時性の範囲は、2音の周波数距離の関数としてV字曲線に回帰することが示されてきた。
- ◆この同時性の範囲のV字型関数は、音圧が低下すると、折点より右側のみで関数の上昇が見られることが、ただひとつの周波数領域条件で報告されていた。
- ◆本研究はこのV字型関数の音圧による変化が高い周波数領域条件でも再現されることを示し、その一般性を示唆した。
- ◆V字の折点より左側の周波数距離では、蝸牛基底膜の振動干渉が寄与していることが考えられる。一方、折点より右側の背後にあるメカニズムは不明である。

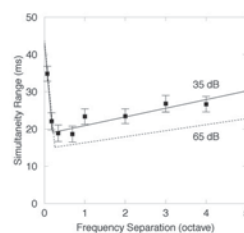


Fig. 1 Perceptual simultaneity range as a function of frequency separation ( $F1 = 200$  Hz).

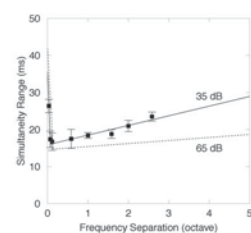


Fig. 2 Perceptual simultaneity range as a function of frequency separation ( $F1 = 800$  Hz).



1-R-6

1-R-6 骨導音併用時の雑音負荷語音聴力

Speech audiometry with combined use of bone-conduction in noisy environments  
 ☆村木悠介, 菅原真一, 森幹男(福井大院)

- ◆骨導音を併用することで、雑音環境下における語音聴力向上効果が期待できるが、その有効性については不明な点が多い。
- ◆雑音環境下で気導音と骨導音同時聴取時の単語理解度試験を行い、気導音と骨導音同時聴取の有効性について検証した。
- ◆雑音環境下では、骨導音を併用して聴取した場合の方が高い単語理解度が得られる傾向があることがわかった。

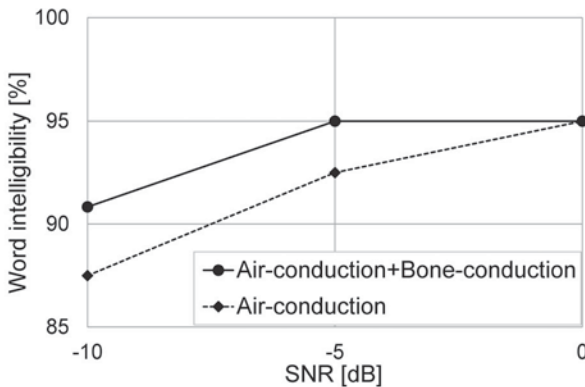


Fig. 1 Result of word intelligibility test in noisy environments

1-R-8

1-R-8 骨導音声の外耳道内放射特性の推定

Estimates of ear-canal radiation characteristics for bone-conducted speech  
 ☆鳥谷輝樹(北陸先端大), △Peter Birkholz(TU Dresden), 鶴木祐史(北陸先端大)

- ◆背景: 骨導音声の伝達過程の解明に向けて、発話器官から側頭部・外耳道までの伝達特性を測定(鳥谷ら, 2019)。  
 →しかし、側頭部・外耳道で得られるそれぞれの伝達特性の関連性を議論するには至らなかった。
- ◆目的: 骨導音声の伝達特性(鳥谷ら, 2019)を用いて、骨導音声の外耳道内放射特性を推定すること。
- ◆方法: 発話器官から側頭部までの伝達関数  $H_b(f)$  と、発話器官から外耳道までの伝達関数  $H_c(f)$  の比により、側頭部振動の外耳道内放射特性  $H_{bc}(f) = H_c(f) / H_b(f)$  を算出

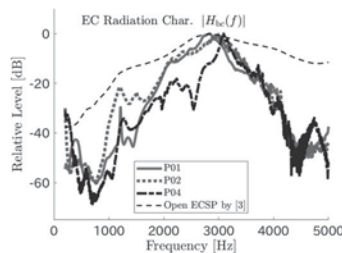


Fig. 1: Estimated ear-canal radiation characteristics

- ◆結果: 3 kHz 付近にピークを有する共通の傾向を確認  
 →頭部振動刺激に対する外耳道内音圧特性(Stenfelt et al., 2003)とピーク位置が一致
- ◆結論: 骨導音声の外耳道内放射特性では3 kHz 付近の成分が強調されることがわかった。  
 この結果は従来の研究結果と一致するものであった。

1-R-7

1-R-7 フォルマント変形聴覚フィードバックにおける補償応答と発話主体感との関係

A study on the relationship between compensatory responses and sense of self-agency under formant transformation auditory feedback  
 ◎上江洲安史, 廣谷定男, 持田岳美 (NTT CS 研)

- ◆フォルマント変形聴覚フィードバック(TAF)とは、話者に特定の音節を繰り返し発話させ、発話音声のフォルマント周波数に摂動を与えた変換音声を話者に実時間フィードバックし、摂動の変化に対する発話の変化(調音補償応答)を調べるものである。
- ◆ヒトが安定した発話を行なうには、フィードバック音声の高域成分(音声自然性)の存在が重要であり、これを外部ノイズで補完できる可能性が示唆されている。このとき、フィードバック音声に対する話者の「自分の声らしさ」すなわち「発話主体感 (SoA)」が変化するために、補償応答にも違いをもたらしている可能性がある。
- ◆本稿では、フィードバック音声の自然性を操作したフォルマント TAF 発話課題と、フィードバック音声に対する SoA 評価課題を並行して行ない、調音補償応答と発話主体感との関係を調べた。

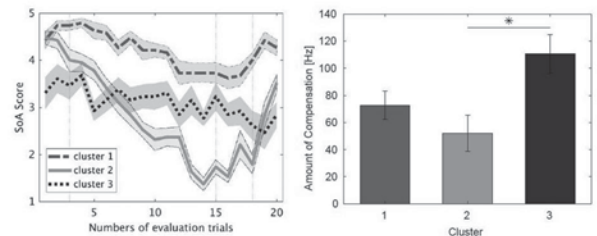


Fig. 1: (Left) Mean patterns of SoA scores clustered by k-means method (Right) Mean values of compensatory responses for each clusters

1-R-9

1-R-9 気導音と骨導音を考慮した音声合成モデルの検討

Voice conversion model based on air-conducted and bone-conducted voices

◎森田翔太, 川本大貴(福山大学・工), 鳥谷輝樹(北陸先端大)

- ◆気導音と骨導音による音声合成モデルを提案する (Fig. 1).
- ◆骨導音は、気導音から推定して求めた。
- ◆音声合成モデルのパラメータ
  - $t_d$ : 骨導音到達から気導音到達に要する時間差
  - $ABR$ : 気導音と骨導音のパワー比
  - $t_r$ : 頭部での反射を仮定した残響の残響時間
- ◆提案した音声合成モデルにより、自己聴取音を再現した (Fig. 2).

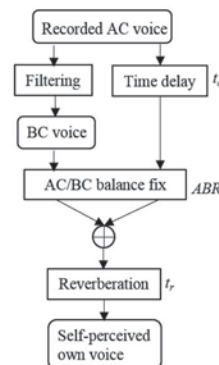


Fig. 1: Voice conversion model.

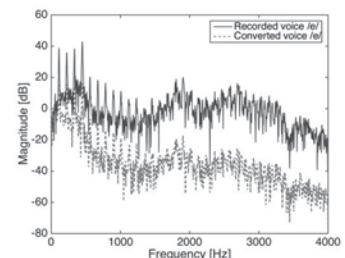


Fig. 2: Spectra of recorded and converted vowels /e/ by a participant. (solid line: recorded voice /e/, broken line: converted voice /e/)

### 1-R-10

#### 1-R-10 音声の危険性の認識に 韻律情報が及ぼす影響 —ストロープ課題を用いた検討—

Effects of prosody information on risk perception of speech on stroop effect

○小林まおり, 赤木正人(JAIST)

- ◆ 音声は言語情報ばかりでなく、パラ言語情報によっても危険性を伝えることができるが、これらが一致していない場合、その判断は困難になることが予想される。しかし、どの程度影響されるのかについては検討されてない。
- ◆ 本研究ではストロープ課題を用いて、音声の危険性の認識に言語情報とパラ言語情報が及ぼす影響について検討した。
- ◆ 実験では、安全性や危険性と関連すると考えられる単語を、安全を意図した話し方と危険を意図した話し方で発話させたものを音声刺激として提示した。被験者は、提示された音声の意味もしくは話し方が、「安全」を意味するか、「危険」を意味するか判断することを求められた。
- ◆ 実験の結果、意味を判断する場合には、単語の意味と話し方が一致する条件では、一致しない条件よりも有意に反応時間が長くなった。一方、話し方を判断する課題では両条件間に有意差は認められなかった。
- ◆ これらの結果から、安全-危険の判断においても、意味と不一致な話し方をすると、判断が困難になることが示された。これは、危険性を伝える避難呼びかけなどにおいても話し方を工夫する必要があることを示唆している。

### 1-R-12

#### 1-R-12 歯を介した骨導音受聴時の 外耳道内音の測定

Measurement of air conduction sound in the ear when listening to bone conduction sound

☆小坂竜位, 森幹男(福井大院)

- ◆ 歯加振時の最小可聴値を主観評価で測定する場合、歯に押力を加え続けることから、被験者には苦痛が伴うという問題点がある。
- ◆ 骨導音受聴時の外耳道内音の音圧を耳内プローブマイクロフォンで測定した。このとき、気導音の回り込みを防ぐために耳栓を着用した。
- ◆ 歯加振時と乳様突起加振時とで比較を行った結果、耳栓の有無によって異なる傾向があることが分かった。

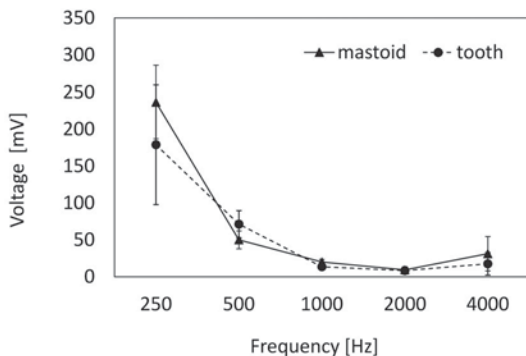


Fig.1 : Sound pressure level comparison using earplugs

### 1-R-11

#### 1-R-11 内側オリブ蝸牛束反射に対する音楽経験 の差：専攻楽器の別と聴力レベルの相関

Effects of musical experience on Medial Olivocochlear Reflex: Difference among speciality and correlations to audiometries.

○津崎 実(京都市立芸大学), 松浦佑美(京都市立芸大学),

大塚 翔(千葉大学, NTT CS研), 古川茂人(NTT CS研),

山本恵利佳(京都市立芸大学)

- ◆ 音楽学部 に所属する学生の内側オリブ蝸牛束反射(MOCR)の程度が専攻する楽器種によって異なるかを調べた。
- ◆ ヴァイオリン専攻生は他の学生よりも大きなMOCRを示した。
- ◆ 先行研究で報告されたような演奏家群と一般群の差は観察されなかった。
- ◆ 聴力レベルとの相関は一部を除いて基本的には見いだされなかった。

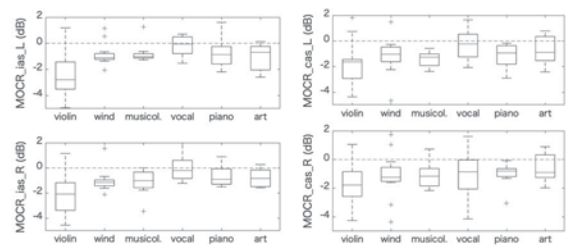


Figure 1. Boxplots of contra-, and ips- lateral MOCRs for left and right ears for the performance specialities.

### 1-R-13

#### 1-R-13 耳栓装用時の骨伝導音知覚/伝搬特性に 及ぼす刺激呈示部位の効果

Effect of the vibrator placement on perception and propagation of bone-conducted sound during earplugging

☆信夫大志, 大塚 翔, 中川誠司(千葉大)

- ◆ 機械工場等の強大騒音下では作業者は耳栓を装着することが多いが、円滑なコミュニケーションが損なわれてしまう。
- ◆ 一方、骨伝導音は耳栓装用下でも明瞭に聴取される。労働者が装着するヘルメット等に骨伝導スピーカを内蔵させることで、強大騒音下でも利用可能な音声コミュニケーション・デバイスが実現できる。
- ◆ 耳栓装用下において、頭頂を含む様々な部位から呈示された骨伝導音の知覚閾や伝搬特性(外耳道内音圧、頭部加速度)を計測した。
- ◆ 頭皮上に呈示した場合は、低域での検出閾の上昇が観察された。
- ◆ 外耳道内放射および頭部加速度の結果から、その主たる要因は外耳道内放射の減少であることが示された。また、圧縮骨導および慣性骨導成分は呈示部位に依存した変化が小さく、むしろ頭皮上の部位に対して増大した可能性が示唆された。

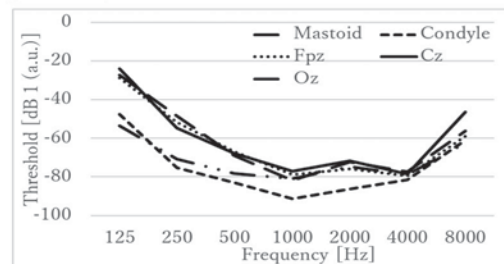


Fig. 1 Sensory threshold at each frequency

### 1-R-14

#### 1-R-14 先行する音列の規則性がオリーブ蝸牛側反射に与える影響

Effect of rhythm-based prediction on medial olivocochlear reflex

○石坂勇毅, 大塚翔, 中川誠司(千葉大)

- ◆オリーブ蝸牛束は、脳幹から外有毛細胞への遠心性の神経であり、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の改善する役割や内耳を強音から保護する機能を果たしていると考えられている。
- ◆刺激音出現のタイミングを知らせる視覚刺激を刺激音に先行して示すことによって、オリーブ蝸牛束反射が強まるということが報告されている。これまでの検討では、単一の先行刺激の影響のみが検討されてきたが、一方で、先行する音列のパターンは次に到来する音を予測するための重要な手がかりになり得る。
- ◆そこで、本研究では、刺激音に先行する音列の規則性から刺激音が到来するタイミングを予測できる場合 (Predictable 条件) と、音列内の音同士の時間間隔がランダムで予測ができない場合 (Unpredictable 条件) でのオリーブ蝸牛束反射の強度を比較した。
- ◆先行する音列が規則的でターゲット音が予測できる場合の方が、オリーブ蝸牛束反射が強く見える傾向が見られた (Fig. 1)。

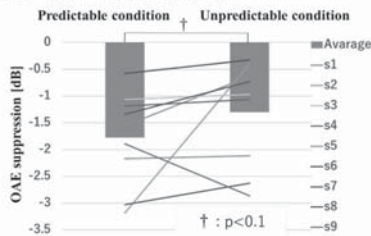


Fig 1: OAE suppression in predictable and unpredictable condition. Each line depicts individual data.

### 1-R-16

#### 1-R-16 聴覚フィルタ推定における蝸牛雑音フロアの設定法について

Cochlear noise floor setting method for auditory filter estimation

☆横田健治, 入野俊夫, (和歌山大),  
Roy D. Patterson (University of Cambridge)

- 聴覚フィルタ推定にレベル依存する蝸牛雑音フロアを導入した。
 
$$N_c^{(LD)}(f) = N_c(dB)(f) + n_{LD} \cdot N_0^{(rel)} + P_s^{(bias)}$$
  - ◇  $N_c(f)$ : 無音状態における蝸牛雑音のスペクトルレベル
  - ◇  $n_{LD} \cdot N_0^{(rel)} (= N_0 - \min(N_0))$ : レベル依存項
  - ◇  $P_s^{(bias)} (= \min(P_s) - P_{abs})$ : バイアス成分
- パラメータ  $P_s^{(bias)}$  を決定する方法の検討
  - ◇ 前回までの発表:  $P_s^{(bias)}$  を固定し、 $N_c(f)$  と比例係数  $n_{LD}$  を収束計算で決定 ← 固定しないと発散する場合があった
  - ◇ 今回2手法を検討 ( $N_c(f)$  は必ず収束計算に入れる)
    1.  $n_{LD}$  を固定値 0.3 か 0.4 に固定 ( $Fix(n_{LD} = 0.3 \text{ or } 0.4)$ )
    2.  $P_s^{(bias)}$  を固定し  $n_{LD}$  を収束計算で決めた後に、 $n_{LD}$  をその値に固定し  $P_s^{(bias)}$  を収束計算で決定 ( $Fix(P_s^{(bias)} \rightarrow n_{LD})$ )
  - ◇ 適合: 信号周波数 1 kHz と 2 kHz の低雑音レベルを含んだノッチ雑音閾値を使用。
- ◆ 推定結果
  - $Fix(P_s^{(bias)} \rightarrow n_{LD})$  において 1 kHz では、従来の設定法よりもノッチ雑音閾値の推定 RMS 誤差と絶対閾値誤差を小さくできた。
  - 2 kHz では、推定 RMS 誤差が従来の設定法より大きくなった。

### 1-R-15

#### 1-R-15 軟骨伝導の出力校正法を目指した聴覚閾および頭部内伝搬の計測

Measurement of auditory threshold and sound propagation in the inner head for development of the calibration method for the cartilage conduction

☆菅原優聖, 大塚 翔, 中川誠司(千葉大)

- ◆骨伝導における振動子の装着感の改善を目的として「軟骨伝導」が提案され、外耳道閉鎖症用の補聴器やスマートフォンに応用されている。
- ◆先行研究においては、軟骨伝導知覚では気導と外耳道内放射が支配的とされているが、呈示条件によっては有意な慣性骨導成分や圧縮骨導成分が発生する可能性が示唆されている。
- ◆軟骨伝導の出力校正法の構築に有用な基礎的データの収集を目的として、軟骨伝導の呈示部位ごとの聴覚閾および頭部内伝搬 (外耳道内音圧、頭部に生じる加速度) を計測した。
- ◆低音域においては、乳様突起と比較して耳介の聴覚閾が低下し (Fig. 1)、耳介と耳前部 (耳珠から顆状突起にかかる部位) での外耳道内音圧が上昇した (Fig. 2)。外耳道内放射成分の効果と考えられる。
- ◆一方、500 - 2000 Hz における耳前部の加速度が他の呈示部位よりも大きかった (Fig. 3)。慣性骨導・圧縮骨導成分の寄与を反映したものと考えられるが、外耳道の軟骨部の加速度を捉えた可能性もある。

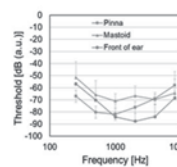


Fig. 1 Threshold

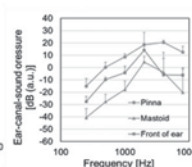


Fig. 2 Ear-canal sound pressure

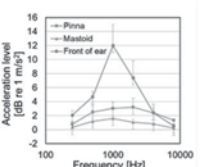


Fig. 3 Acceleration of the head

### 1-R-17

#### 1-R-17 聴覚フィルタを用いた正中面頭部伝達関数の帯域間相関

Analysis of inter-band correlation in HRTF in median plane with auditory filter

○岩谷幸雄(東北学院大学/Sorbonne 大学),

Brian FG Katz (Sorbonne 大学), 曲谷地哲, 鈴木陽一(東北大学)

- ◆目的: 正中面内の音像定位における物理的なスペクトラルキューと、聴覚処理におけるスペクトルの大局的な特徴を結びつける。
- ◆実施事項: 東北大学で測定された SOFA 準拠の 105 セットの頭部伝達関数について、正中面内の頭部伝達関数を 40 帯域の ERB に分割し、それらの角度に対するレベル変化を求め、帯域間の相関を求めた。
- ◆結果: 共通的な傾向として、ノッチが系統的に周波数シフトする 6 kHz 付近の帯域が、他の帯域に対して高い負の相関を得ることが分かった。1 kHz 付近における帯域も負の相関をとる傾向にあるが、実際のレベル変化が小さいため、音像定位への影響は小さいと予想される。

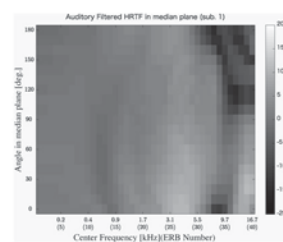


Fig.1: An example of filtered HRTF divided into 40 ERBs in upper median plane

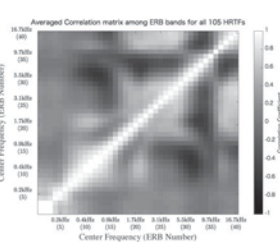


Fig.2: Correlation matrix of inter-ERB averaged across all 150 sets of HRTF