

**科学研究費助成事業 研究成果報告書**

平成 27 年 5 月 25 日現在

機関番号：25406

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24560529

研究課題名(和文) 騒音環境下での音声認識を目指した体内伝導音計測に基づく音声信号抽出法

研究課題名(英文) Detection method of speech signal based on the measurement of body conducted sound for speech recognition under noise environment

研究代表者

生田 顯 (Ikuta, Akira)

県立広島大学・経営情報学部・教授

研究者番号：30145164

交付決定額(研究期間全体)：(直接経費) 4,100,000円

研究成果の概要(和文)：ハンズフリー業務等の支援に音声認識を応用することを目的とし、騒音環境下での音声信号抽出法を開発した。具体的には、騒音混入下での観測に基づき、音声信号のみを推定するための信号処理法を、信号や観測値の有限振幅変動に整合したベータ分布の導入により提案した。次に、騒音環境下での気導音と体内伝導音を同時計測し、鮮明な音声信号を抽出するための適応ノイズキャンセラー(ANC)を開発した。具体的には、FLANNおよびVolterra filterを用いた方法、さらに両者を併用したハイブリッドANCシステムを提案した。提案手法の有効性は、無響室で測定した、騒音下での音声と体内伝導音に適用することにより確認した。

研究成果の概要(英文)：For applying speech recognition to actual circumstances where hand-writing is difficult, a detection method of speech signal was considered under existence of surrounding noise. More specifically, a signal processing method to estimate only the speech signal based on the observation data contaminated by noise was proposed by introducing Beta distribution appropriate for the finite amplitude fluctuation range of signal and observation. Next, both body- and air-conducted speech signals were used simultaneously to fight up against the surrounding noise to improve the enhancement quality. Nonlinear adaptive noise canceller (ANC) was proposed by using FLANN and Volterra filter. Furthermore, a hybrid system combining both FLANN and Volterra filter was proposed. The effectiveness of the proposed system was confirmed by applying it to real body- and air-conducted speech signals measured in an anechoic chamber under existence of surrounding noise.

研究分野：音響信号処理

キーワード：音声信号 雑音除去 体内伝導音 気導音 実環境 音声認識

## 1. 研究開始当初の背景

(1) デジタル情報技術の進展に伴い、音声認識システムが開発され、様々な分野に応用されているが、この分野においては、周囲の騒音が存在する実環境下での騒音対策が、現在、最も重要な研究課題である。

研究代表者は、これまで音響システムにおける雑音除去や適応信号処理に基づく能動騒音制御に関する研究を行っていた。さらに、騒音の混入した音声データから音声信号のみを推定するための確率的信号処理法を提案した。従来多くの雑音除去法は、変動雑音のガウス分布性・白色性を前提としているのに対し、上記の研究では非ガウス・非白色性の雑音に対しても有効であり、実環境における状況により整合しているといえる。

本研究では、実環境下での音声認識を可能にするため、周囲の騒音が入りにくい固体伝播音である体内伝導音の計測に基づき、音声信号を抽出するための新たな処理技術を研究開発した。

(2) 従来から、音声認識における騒音対策法はいくつか提案されてきたが、それらの多くは、騒音の存在下においてマイクロフォンで測定された音声信号に対し、何らかの信号処理を施すことにより騒音の影響を抑圧するものである。

一方、体内伝導音は、声帯で発生した音声身体表面へ伝播する振動である固体伝播音であるため周囲の騒音の影響を受けにくい。したがって、騒音環境下における音声認識が可能である。しかし、加速度ピックアップにより身体表面で計測された体内伝導音は、マイクロフォンにより測定された音声と比べて高周波成分の減衰が大きく、情報欠損を生じる。また、心音などの体内雑音の影響を受ける。

改善法がいくつか提案されているが、それらの手法では体内伝播を表す音声モデルを事前に推定する必要がある。これらの手法の最大の問題点は、音声モデルパラメータを騒音が存在する現実環境下で事前に推定することが困難なことにある。また、雑音の抑制法としてウィナーフィルタやLMS(最小二乗法)に基づく適応フィルタを適用した手法が提案されている。しかし、これらのフィルタは雑音の変動がガウス分布に従うことや白色性を有することを前提としており、しかも信号の線形相関情報のみを利用している。しかし多くの雑音の変動は実際には非ガウス型の複雑な形態を示し一般には非白色で、現実の信号には高次の相関情報も存在しているなど現実問題との大きな乖離があり、有効な音声抽出法は見出されていなかった。

## 2. 研究の目的

(1) 現実の騒音環境下での音声認識を可能にするため、音声モデルが未知の実際状況を対象とした。すなわち、音声モデルパラメータを音声信号から事前に推定する必要がなく、音声信号や体内伝導音、体内雑音に対する変動のガウス特性に対する前提を必要としない新たな音声信号抽出法の開発を目的とした。

体内伝導音に対する線形あるいは非線形の音声モデルパラメータを推定するためには、一般に音声信号と体内伝導音の相関情報を必要とする。しかし、音声信号は未知の推定対象であるため、その情報を事前に把握することは一般に困難である。また、音声信号は複雑な変動形態を示すのが通常であり、音声信号に対する高精度の推定を目指すためには、音声信号と体内伝導音の揺らぎ形態の全情報を反映した結合確率分布全体に着目する必要がある。

本研究の目的は、音声信号と体内伝導音の関係として、非線形の音声時系列モデルを導入することにより、新たな音声信号抽出アルゴリズムを提案することである。

(2) 騒音環境下での音声観測値(気導音)と体内伝導音を同時計測し、鮮明な音声信号を抽出するため、新たな適応ノイズキャンセラー(ANC)の研究開発を目的とした。すなわち、気導音をプライマリー信号、体内伝導音を参照信号とした線形および非線形の適応ノイズキャンセラーについて考察した。

(3) 提案した手法の有効性を実際に測定した騒音環境下での音声信号や体内伝導音に適用することにより、実験的に検証することを最終的な目的とした。

## 3. 研究の方法

(1) 体内伝導音の計測に基づく音声信号抽出アルゴリズムの開発

工場など生産の現場における施設・設備巡回点検修理業務や工事現場の記録・状況報告などハンズフリー業務の支援に音声認識を応用することを目的とし、周囲環境からの騒音存在下での音声信号抽出技術に関し新たな実用的手法を開発した。具体的には、周囲騒音の影響を受けにくい体内伝導音の計測に基づき音声信号を抽出するための信号処理法について検討を行った。本研究では、実環境下での音声認識を可能にするため、雑音変動のガウス性や白色性などの前提を必要とせず、しかも信号に関する高次の非線形相関情報をも考慮し、かつ音声の時間的変動形態が未知の場合でも適用できる新たな音声信号抽出法の研究開発を目指した。

(2) 開発した音声信号抽出アルゴリズムの有効性確認

騒音存在下での実環境で有効な音声認識システムを構築するために、音声認識システムの応用が期待されている実際の現場における音声データおよび体内伝導音を測定・分析した。特に、騒音下でのアルゴリズムの適用可能範囲について、様々な SN 比において無響室で測定した実データをもとに詳細に検討を行った。音声認識率のさらなる向上や工場等の生産の現場への応用が可能となる音声認識システムを構築し実用化を計るために、研究開発した体内伝導音計測に基づく音声信号抽出法を、実際の騒音下における音声データに適用し、その実際的な有効性を確認した。

さらに、実環境下での音声認識を可能にすべく、開発した音声に関する非線形時系列モデルをもとに、システムの精密化を図った。

#### 4. 研究成果

(1) 実環境での計測においては、振幅変動の有限性や計測器のダイナミックレンジにも基づいて、現実に撮取できる信頼性のある振幅レベルの変動範囲としては、有限のレベル範囲内に限定された情報であるのが通常であり、また、分解能を上げて精密な測定を行う場合は、レコーダのダイナミックレンジをあえて狭く設定することもしばしばであることに留意した。すなわち、騒音混入下での音声信号に対する観測データに基づき、音声信号のみを推定するための信号処理法を、信号や観測値の有限振幅変動に整合したベータ分布を重みとする直交展開型分布表現を導入することにより提案した。

提案した騒音除去法の有効性を確認するため、広島県立総合技術研究所西部工業技術センターの音響実験棟における無響室で測定した騒音下での音声データに本手法を適用した。具体的には、白色雑音下での女声信号および有色雑音（ピンク雑音）下での男声信号それぞれに対する観測値（サンプリング周波数 10 kHz、量子化 16 bits）に基づき音声信号の逐次推定を行った。

各観測波形を、それぞれ Fig. 1 および Fig. 2 に示す。本研究で開発した手法を適用することにより音声信号を推定した結果を、それぞれ Fig. 3 および Fig. 4 に示す。比較のため、信号や観測値の有限振幅変動を考慮せずに、ガウス分布を重みとする直交展開型分布表現に基づく推定アルゴリズムを用いた女声および男声信号の推定結果を、それぞれ Fig. 5 および Fig. 6 に示す。

女声信号については Fig. 3 と Fig. 5 を、男声信号については Fig. 4 と Fig. 6 を比較することにより、音声信号の有限変動域を考慮

した推定アルゴリズムを用いた結果に比べて、音声信号の変動域を  $(-\infty, \infty)$  に仮定したガウス分布を基幹とする推定アルゴリズムを用いた場合は、雑音の影響が十分に除去できていないことがわかる。

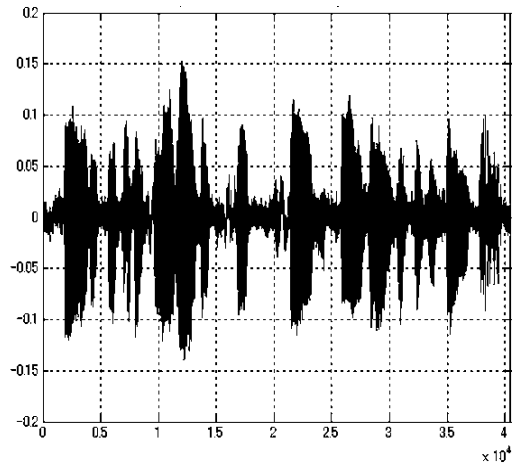


Fig. 1 白色雑音下での女声信号

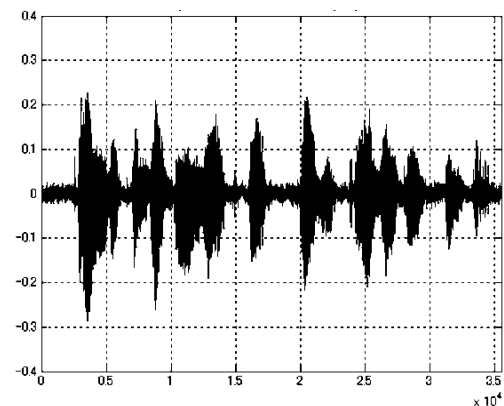


Fig. 2 有色雑音下での男声信号

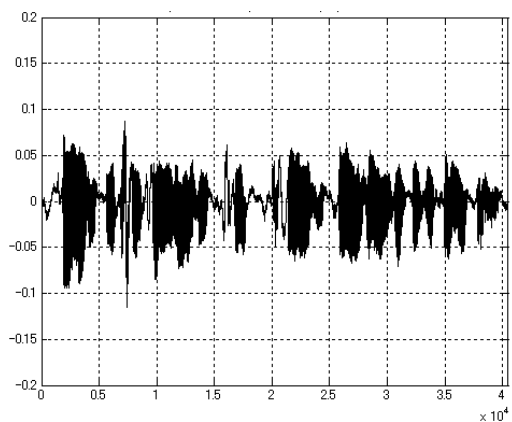


Fig. 3 女声信号の推定結果

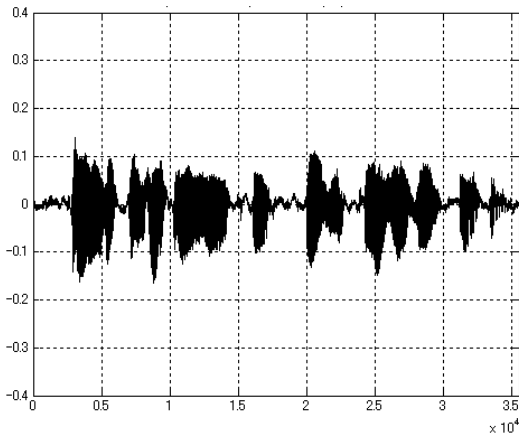


Fig. 4 男声信号の推定結果

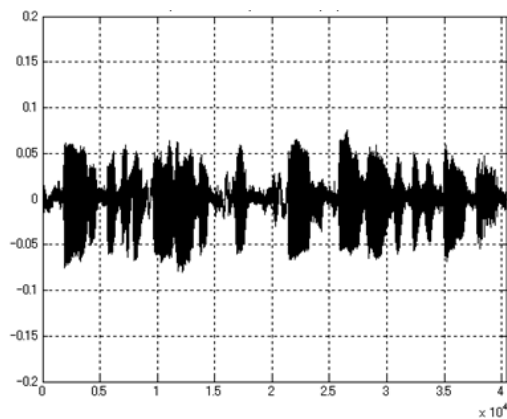


Fig. 5 ガウス分布展開に基づく女声信号の推定結果

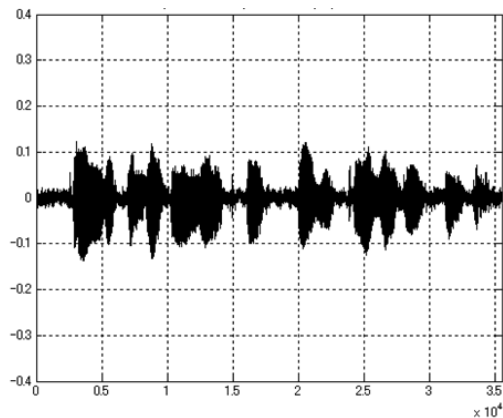


Fig. 6 ガウス分布展開に基づく男声信号の推定結果

(2) 実環境下での音声認識を可能にするために、マイクロフォンで測定される周囲環境騒音が混入した気導音と、騒音の影響を受けにくい体内伝導音を同時に用いた新たな適応ノイズキャンセラー(ANC)を開発した。すなわち、気導音をプライマリー信号、体内伝導音を参照信号とし、最小二乗規範(LMS)による線形適応ノイズキャン

セラーを提案した。この手法では、騒音をかなり抑制し音質を大きく改善するが、音声信号の高周波成分の復元には不十分であることが実験により確認された。その理由は、体内伝導音から気導音までの物理的なシステム特性には非線形性が含まれていることに起因する。このような観点から、非線形適応フィルタとして学習が容易で性能も良好とされる Functional Link Artificial Neural Network(FLANN)を用いた、騒音除去のための新たな信号処理法を提案した。従来の線形 ANC および提案の非線形 ANC をそれぞれ Fig. 7 および Fig. 8 に示す。

提案法の有効性を検証するため、無響室において同期を取りながらマイクロフォンで気導音を、振動ピックアップで体内伝導音を計測した。気導音に雑音を加えた信号をプライマリー信号とした。まず、線形 ANC を応用してみた。気導音に加える雑音の分散を小さい値から徐々に大きくしてシミュレーション実験を行った。その結果、線形 ANC でも復元する能力があることが分かった。ただし、用いる FIR フィルタの長さを大きくする必要があり、本実験では  $L=128$  に設定した。

次に、提案システムを用いて、音声復元を行った。特に雑音が多い場合について、様々なパラメータ設定でシミュレーション実験を行った。ここに代表的な結果を示す。Fig. 9 には、上から、雑音のない気導音(原音声)と雑音に乱された気導音(雑音の分散が原音声のそれと同じ値になるよう雑音の分散を調整した)、体内伝導音、線形 ANC(Fig. 7)により復元された音声信号、本提案の非線形 ANC(Fig. 8)により復元された音声信号(男性)をそれぞれ示す。

実験結果から、次のことが分かった。1) 線形 ANC も提案した非線形 ANC も、ひどく汚された観測信号から原音声を抽出する能力を有する。2) 本提案法は、線形 ANC に比べて、高周波数成分の回復においてより高い復元能力を有する。

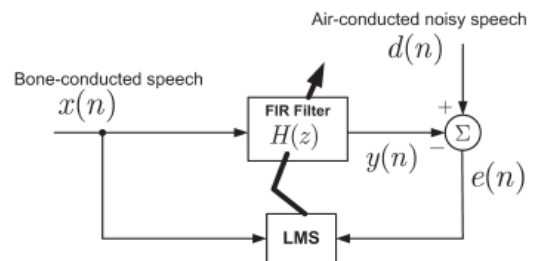


Fig. 7 線形 ANC

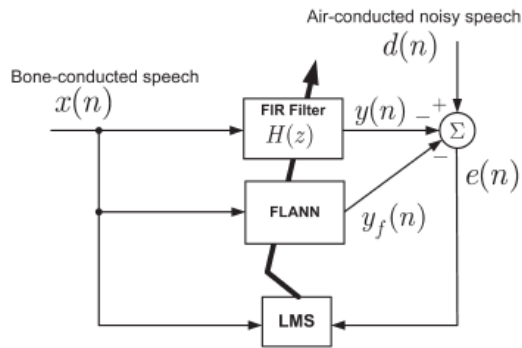


Fig. 8 FLANN による非線形 ANC

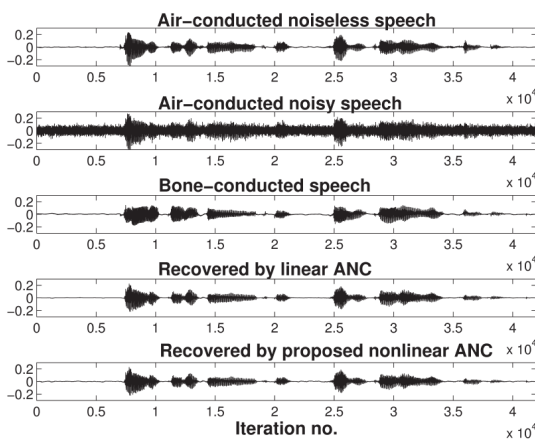


Fig. 9 音声信号に対する復元結果の比較

(3)(2)で提案した手法は Neural Network のサイズや重みの初期値設定など、試行錯誤を要し処理も複雑であった。このような観点から、次に、非線形適応フィルタとして広く用いられている Volterra filter を用いた方法を提案した。図示により FIR と Volterra filter を比較することが困難なため、復元音声と原音声との平均自乗誤差(MSE)を Table 1 にまとめた。

実験結果から、本提案法は、線形 ANC に比べて、誤差の自乗平均においてより高い復元能力を有することが明らかになった。

Table 1 FIR filter と Volterra filter による平均自乗誤差(MSE)の比較

(Ma: Male, Fe: Female, Vol: Volterra)

		MSE ( $\times 10^{-4}$ )					
SNR (dB)		100	10	1	0	-1	-3
Ma	FIR	1.49	1.51	1.69	1.73	1.80	1.98
	Vol	1.18	1.21	1.45	1.52	1.60	1.85
Fe	FIR	1.89	1.89	1.91	1.92	1.93	1.95
	Vol	1.57	1.57	1.59	1.60	1.61	1.63

(4)FLANN と Volterra filter の両者を併用したハイブリッド ANC システムを提案した。復元音声と原音声との平均自乗誤差(MSE)について、4種類の ANC を適用した場合を Table 2 に示す。FLANN がべき乗の非線形近似に適しているのに対し、Volterra filter は変数間の乗算項に対応した非線形関数の近似に適していることから、両者の併用は互いの長所・短所を補完し合い、単独使用より優れた性能を示すことが実験的にも確認できた。

Table 2 4種類の ANC を適用した場合の復元音声と原音声との平均自乗誤差 (MSE)

(Ma: Male, Fe: Female, Vol: Volterra, FLA: FLANN, Pro: Proposed)

		MSE ( $\times 10^{-4}$ )					
SNR (dB)		100	10	1	0	-1	-3
Ma	FIR	1.18	1.20	1.32	1.36	1.41	1.54
	Vol	1.07	1.09	1.24	1.29	1.35	1.51
	FLA	1.07	1.09	1.25	1.29	1.35	1.51
	Pro	0.98	1.01	1.19	1.24	1.32	1.50
Fe	FIR	1.32	1.32	1.33	1.34	1.34	1.36
	Vol	1.24	1.24	1.26	1.27	1.27	1.29
	FLA	1.07	1.07	1.10	1.11	1.11	1.15
	Pro	1.02	1.02	1.05	1.06	1.06	1.10

## 5. 主な発表論文等

[雑誌論文](計7件)

A. Ikuta, N. H. Siddique and H. Orimoto: Prediction of Output Response Probability of Sound Environment System Using Simplified Model with Stochastic Regression and Fuzzy Inference, Fuzzy Information and Engineering, 査読有, Vol. 11, No. 2, pp. 173-190 (2013).

A. Ikuta and N. H. Siddique: A Cancellation Method of Additive External Noise and State Dependent Noise in Sound Environment Systems, International Journal of Applied Pattern Recognition, 査読有, Vol. 1, No. 3, pp.224-241 (2014).

A. Ikuta and H. Orimoto: Stochastic Approach for Noise Suppression of Speech Signal by Considering Finite Range of Amplitude Fluctuation in Real Environment, International Journal of Applied Pattern Recognition, 査読有, Vol. 1, No. 4, pp.361-376 (2014).

H. Orimoto and A. Ikuta: State Estimation Method of Sound Environment System with Multiplicative and Additive Noises, International Journal of Circuits, Systems and Signal Processing, 査読有, Vol. 8, pp. 307-312 (2014).

A. Ikuta and H. Orimoto: Stochastic Signal Processing for Sound Environment System with Decibel Evaluation and Energy Observation, *Mathematical Problems in Engineering*, 査読有, Vol. 2014, Article ID 208612, 8 pages, <http://dx.doi.org/10.1155/2014/208612> (2014).

A. Ikuta, H. Ogawa and H. Orimoto: Signal Processing in Power State Variables for Noise Cancellation of Electromagnetic Environment, *Electrical Engineering in Japan*, 査読無, 学会からの依頼, Vol. 189, No. 1, pp. 36-46, (2014).

A. Ikuta, H. Orimoto, N. H. Siddique and L. P. Maguire: Statistical Evaluation of Complex Sound Environment with Background Noise, *International Journal of Acoustics and Vibration*, 査読有, Vol. 20, No. 1, pp. 41-46 (2015).

[国際会議](計8件)

A. Ikuta and H. Orimoto: Stochastic Signal Processing for Noise Suppression of Speech Signal in Real Environment, *Proceedings of 19th International Congress on Sound and Vibration*, 査読有, 8 pages, (July, 2012).

A. Ikuta and Ran Xiao: Stochastic Approach for Adaptive Noise Suppression of Speech Signal in Real Environment, *Proceedings of the 11th IEEE International Congress on Cybernetic Intelligent Systems*, 査読有, pp.81-86, (August, 2012).

A. Ikuta and S. Arima: Cancellation Method of Additive External Noise and State Dependent Noise in Sound Environment Systems by Use of Stochastic Signal Processing, *Proceedings of 20th International Congress on Sound and Vibration*, 査読有, 8 pages, (July, 2013).

R. Xiao and A. Ikuta: A Speech Signal Processing Method by Considering Finite Range of Amplitude Fluctuation and Its Application to Actual Environment, *Proceedings of 20th International Congress on Sound and Vibration*, 査読有, 8 pages, (July, 2013).

A. Ikuta and H. Orimoto: Stochastic Signal Processing for Cancellation of Additive and Multiplicative Noises in Sound Environment Systems, *Proceedings of Inter Noise 2013*, 査読有, 10 pages, (September, 2013), (Invited Paper).

A. Ikuta and H. Orimoto: Noise Suppression of Speech Signal by Considering Finite Range of Amplitude Fluctuation in Real

Environment, *Proceedings of Inter Noise 2013*, 査読有, 8 pages, (September, 2013).

A. Ikuta and H. Orimoto: Signal Processing in Power State Variables for Sound and Electromagnetic Environment by Introducing Fuzzy Probability, *Proceedings of the 2014 International Conference on Communications, Signal Processing and Computers*, 査読有, pp.19-26, (February, 2014), (Best Papers).

A. Ikuta and H. Orimoto: Fuzzy Signal Processing of Sound and Electromagnetic Environment by Introducing Probability Measure of Fuzzy Events, *Proceedings of International Conference on Fuzzy Computation Theory and Applications*, 査読有, pp. 5-13 (October, 2014), Best Paper Award (最優秀論文賞受賞)

[学会発表](計5件)

K. Doi, B. Huang, Y. Xiao and A. Ikuta: Narrowband ANC Using Variable Step-size FXLMS Algorithm, *電子情報通信学会技術研究報告(応用音響)*, pp.147-152, (2012. 11)

生田 顕, 肖 然, 長谷川浩治, 太田光雄: 振幅変動の有限性を考慮した音声信号処理法と実環境への適用, *電子情報通信学会技術研究報告(応用音響)*, pp.153-158, (2012. 11)

生田 顕, 肖 業貴, 折本寿子, 長谷川浩治: 骨導音と気導音を用いた非線形適応ノイズキャンセラーによる音声復元, *日本音響学会講演論文集*, pp.681-684, (2014. 3)

肖 業貴, 肖 然, 生田 顕, 折本寿子, 長谷川浩治: Volterra filter を用いた適応ノイズキャンセラーによる音声復元, *日本音響学会講演論文集*, pp.557-560 (2014. 9)

肖 然, 肖 業貴, 生田 顕, 折本寿子, 長谷川浩治: FLANN と Volterra filter を併用した適応ノイズキャンセラーによる音声復元, *日本音響学会講演論文集*, pp. 1019-1022 (2015. 3)

## 6. 研究組織

### (1) 研究代表者

生田 顕 (IKUTA AKIRA)

県立広島大学・経営情報学部・教授

研究者番号: 30145164

### (2) 研究分担者

折本寿子 (ORIMOTO HISAKO)

県立広島大学・経営情報学部・講師

研究者番号: 80533207