



テーマ名	時変性がある雑音環境下でもクリアな音声を抽出することが可能な音声処理技術
組織名	埼玉大学 理工学研究科 島村 徹也 教授
技術分野	IT

### 概要

日常環境下では様々な雑音が存在します。定常的な雑音については除去技術が進んでいますが、時間変化が大きい雑音の除去は困難でした。本研究室では、時変性がある雑音に対しても有効に働くフレーム内処理方法等の新しい雑音低減・除去技術を持っています。これは、電話等の通話品質だけでなく、音声による個人認証や機器のコントロールにも有効であり、また、音声に限らず画像等にも適用可能です。また、深層学習の利用にも対応しています。本技術を用いた応用製品の開発や事業展開に意欲がある企業を歓迎します。

### 簡略図

時変性がある雑音環境下でもクリアな音声を抽出することが可能な音声処理技術

現フレームのみを用いる雑音抑圧処理技術

#### 【特徴】

- ・リアルタイム処理で、フレーム処理ベースの各種雑音低減手法に利用可能
- ・時変雑音(時間変化の大きな雑音)を含む各種騒音環境下において雑音抑制
- ・音声に限らず画像等にも適用可能

#### 【応用先】

【応用例】 通話品質向上だけでなく音声認識への応用も可能



操作の指示を音声で認識



騒音環境下の音声通話



話者認識を用いたセキュリティサービス



### 背景

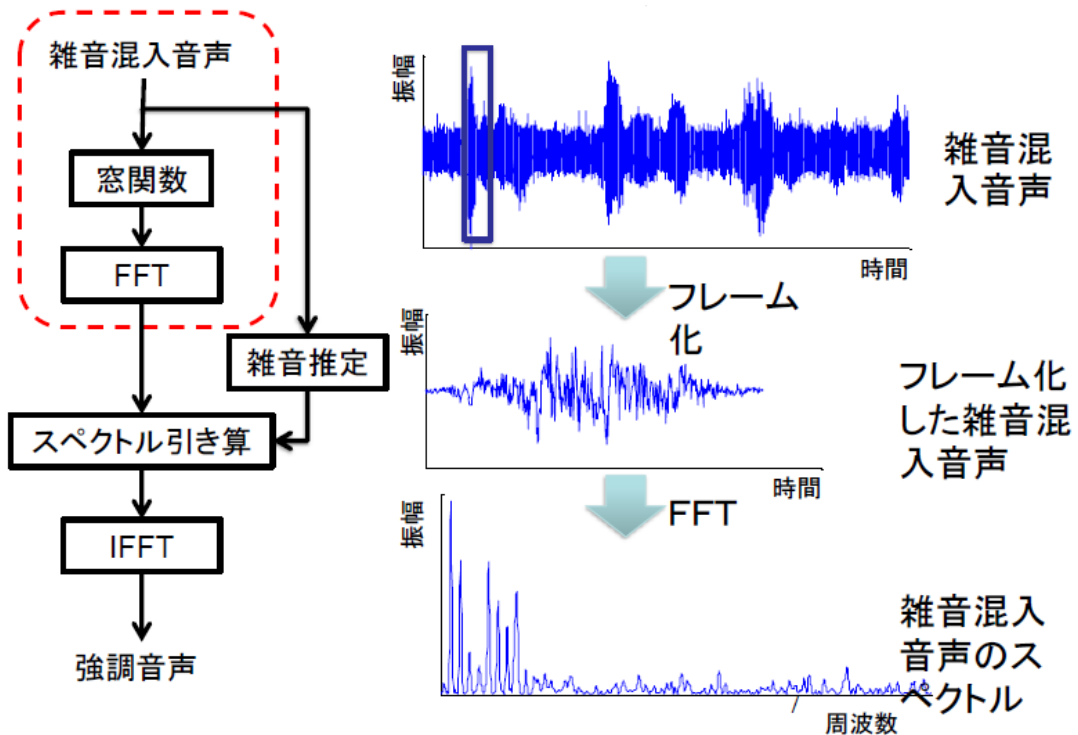
日常環境下では様々な雑音が存在し円滑なコミュニケーションの障害となっています。例えば、電話の通話等で周辺環境がうるさいと音声信号に雑音が混ざってしまい、聞き取りにくいという結果になります。この他にもクリアな音声は音声による個人認証や機器のコントロールにも必要です。



雑音環境下では音声品質が劣化

高性能な雑音抑圧処理が必要

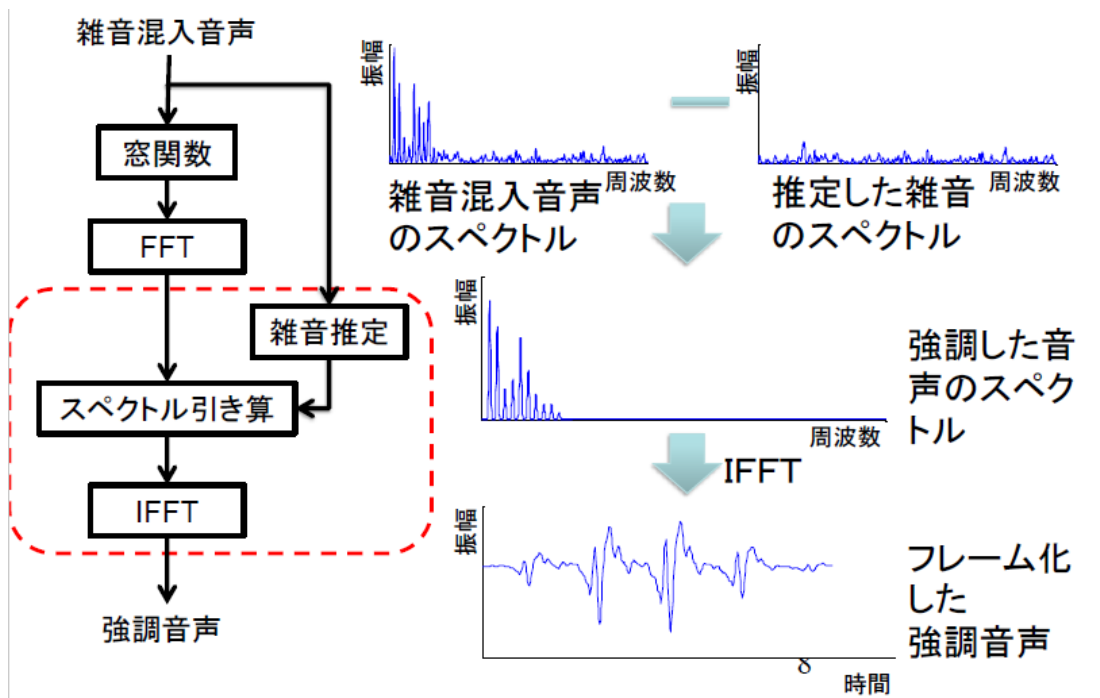
従来、雑音低減・除去技術として良く知られているのが、スペクトル引き算です。



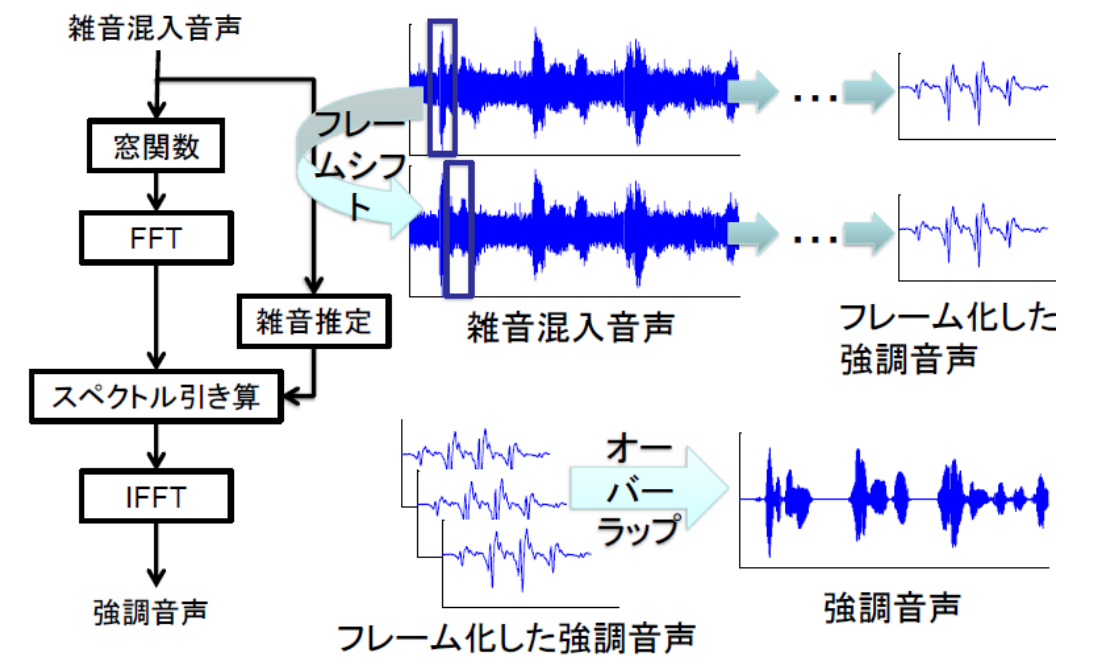
この方法では、雑音の混入した音声を、まず、時間幅数十 $\mu$ 秒程度のフレームに区切



った後、高速フーリエ変換(FFT)します。この状態では上図右下のように、雑音と音声  
が混ざったスペクトルとなっています。一方で雑音推定を行います。すなわち、雑音の  
スペクトルを推定します。そして、下図のように雑音と音声混ざったスペクトルから  
雑音のスペクトルを差し引き(スペクトル引き算)し、残った音声スペクトルを逆変換  
(IFFT)して、音声信号を取り出します。



フレームをシフトしながらこの操作を繰り返し、結果をオーバーラップさせることで雑音を低減・除去した強調音声を得られます。





しかし、この方法では「雑音推定」が本質的に重要です。雑音推定方法としては、例えば、無音声区間を用いた雑音推定法があります。これは無音声区間では雑音のみが存在することを利用した方法ですが、「雑音は無音声区間に流れている雑音と常に同じである」ことが前提になります。すなわち、非定常的な雑音には対応できません。

過去のフレームを用いて統計的手法で雑音推定する方法も提案されています。例えば、過去のフレームの最小統計量で雑音推定する最小統計法があります。これは非定常的な雑音にも対応するものの推定遅延が生じるという問題があります。

このように、「現在のフレーム」のみで非定常的な雑音についても雑音推定を行うことが望ましく、この方法としては松川法と呼ばれる方法があります。これは、非定常の雑音に対応可能ですが、白色雑音のような雑音にのみ有効です。

なお、雑音の低減・除去方法としては、上記のスペクトル引き算法のほか、

- ・ウィナーフィルタ
- ・くし形フィルタ
- ・適応フィルタ(adaptive filter)の利用 (カルマンフィルタを含む)
- ・順序統計フィルタ
- ・各種非線形フィルタ
- ・ノッチフィルタ
- ・ニューラルネットワーク

等の手法が知られていますが、「現在のフレーム」のみで非定常的な雑音について低減や除去を行うことは困難でした。

### 技術内容

本研究室で開発した雑音低減・除去技術は、以下の特徴を有します。

- ・各種騒音環境下において、効果的に音声を強調 (SN 比向上) できる。
- ・実環境によくある時変雑音を、追従することなしに、雑音低減可能。
- ・リアルタイム処理で、フレーム処理ベースの各種雑音低減手法に利用できる。

このため、時変性がある雑音環境下でもクリアな音声を抽出することが可能であり、電話等の通話品質だけでなく、実環境下での音声による制御、認識、管理などに適用できます。すなわち、従来の雑音低減・除去技術では困難であった音声による個人認証や機器のコントロールにも適用でき、音声に限らず画像におけるノイズの除去等にも適用可能です。また、画像のみならず音声についても有意義なニューラルラーニングを適用することが可能です。

### 技術・ノウハウの強み(新規性、優位性、有用性)

本研究室で開発した雑音低減・除去技術の利点は以下の通りです。

- ・現フレームのみを用いる雑音抑圧処理法である。
- ・リアルタイム処理で、フレーム処理ベースの各種雑音低減手法に利用できる。
- ・複数の手法を有しており、実環境によくある時変雑音 (時間変化の大きな雑音) を含



む各種騒音環境下において、効果的に音声を強調（SN比向上）できる。

- ・ミュージカルノイズ（残留雑音）や音声スペクトルの歪みが少ない。
- ・ニューラルネットワークの適用も可能である。
- ・音声に限らず画像等にも適用可能である。

### 連携企業のイメージ

例えば下記の企業と連携可能です。

- 1) モバイル端末やその応用に関わる企業
- 2) 自動車の自動運転における音声指示や家電・AI機器の音声操作、音声応用分野への展開を考えている企業
- 3) 音声によるセキュリティーシステムに関わる企業
- 4) 音声に関してニューラルネットワークの適用に興味のある企業
- 5) その他、各種状況下（例えば、工場、工事現場等の高騒音環境や、ドローン等の騒音源近くでの作業）での雑音の除去を望んでいる企業
- 6) 画像鮮明化など、音声以外への本技術に適用に関心のある企業
- 7) その他、本技術の製品化・活用に意欲がある企業。

### 技術・ノウハウの活用シーン(イメージ)

このように、本研究室で開発した雑音低減・除去技術は、実生活環境のような騒音が大きな環境下でクリアな音声を抽出することに適しています。



従って、実環境下での音声通話の品質改善に役立つのはもちろん、音声認識性能の改善にも有効であり、

- ・携帯端末等での音声認識
- ・家庭内での電化製品・AI機器等への音を用いた指示や会話
- ・実環境での音声の書き出し（テキスト化）

などの用途を有するほか、より厳密な音声認識が必要な

- ・車の自動運転に向けての音声認識や話者認識
- ・音によるセキュリティーシステム

にも応用が可能です。



操作の指示を  
音声で認識

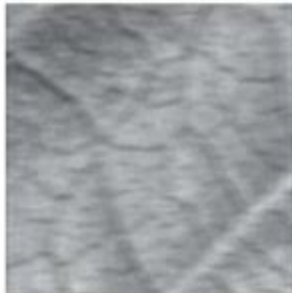


騒音環境下の  
音声通話

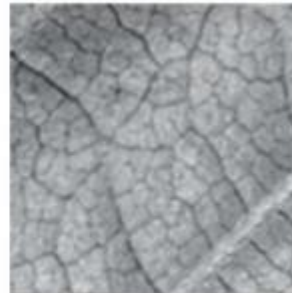


話者認識を用いた  
セキュリティサービス

また、本研究室における雑音低減・除去技術は、以上のほか、信号と雑音を分離したい各種分野（海洋、人体、生体、音楽など）に適用可能なほか、画像への適用も可能です。下記はその一例です。



〈ノイズ混入ボケ画像〉



〈復元画像〉

### 技術・ノウハウの活用の流れ

本技術の活用や製品開発に興味がある方はお気軽にお問合せください。  
デモを交えてご紹介させていただきます。

### 専門用語の解説

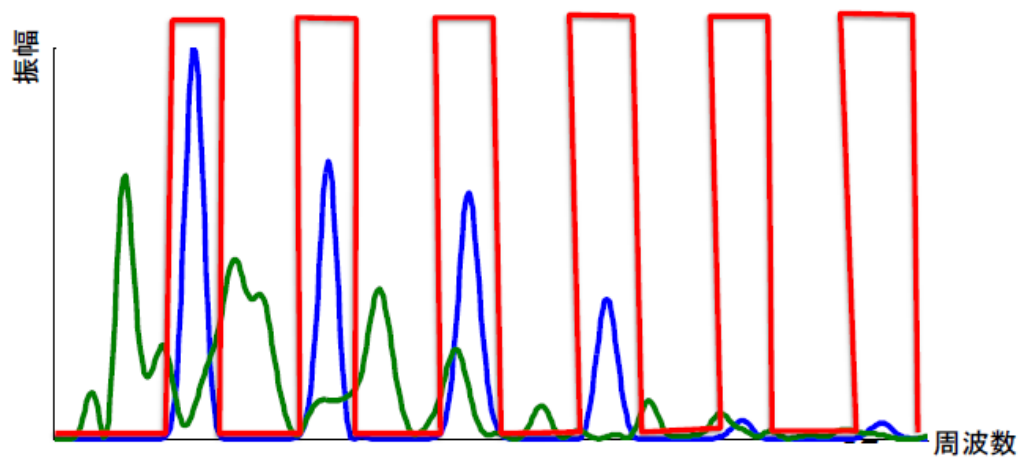
#### 【高速フーリエ変換(FFT: fast Fourier transform)】

デジタル信号の周波数解析で使われる離散フーリエ変換をコンピュータ上で高速に計算するアルゴリズムです。逆変換を逆高速フーリエ変換（英: inverse fast Fourier transform, IFFT）と呼びます。



【くし形フィルタ】

周波数特性が一定間隔のスパイク状（下図の赤線）になったフィルタです。フィルタリングイメージが「くし(comb)」状であることからくし形フィルタと呼ばれます。音声スペクトルが基本周波数の整数倍に存在することを利用して音声（下図の青線。緑は雑音の例）の抽出に用いられることもあります。カラーテレビで輝度信号と色信号の分離をするためのフィルタとしても有名です。



お問合せ先

下記から御問合せください。

<http://www.open-innovation-portal.com/university/it/saitama.html>