

100dB の騒音下でも高品質な通話や音声認識を可能とする小型インテリジェントマイクを開発

～様々な利用シーン向けの音響信号処理技術を取り揃え、NTT-AT より展開～

日本電信電話株式会社(以下 NTT、本社:東京都千代田区、代表取締役社長:鶴浦博夫)は、100dB の騒音下(工場内や電車通過中のガード下など)でも、高品質な通話や高精度な音声認識を可能とする「小型インテリジェントマイク」を開発しました。

小型インテリジェントマイクとは、安価な汎用のマイク素子 2～3 個と、目的とする音声とそれ以外の雑音を高精度に推定して分離する「音響信号処理技術」を組み合わせることで、高騒音環境下でも、話者の音声を明瞭に集音できるものです。(図 1)

本マイクを活用することで、これまで実現が困難であった大きな騒音下、例えば、工場内や工事現場等の機械騒音下や、高速道路を走行中の自動車内などにおいても、高品質な通話(明瞭なハンズフリー通話)や高精度な音声認識が可能になります。

これまで、NTT では、“単一マイク”で構成されるスマートフォンや音声会議装置、録音機器でも、高品質な通話・録音、音声認識を可能とする「背景雑音抑圧技術」※や、カメラでズームするように遠く離れた場所(約 20m)で話す人々の声から指定した人の声のみをクリアに集音可能な「ズームアップマイク技術」*1 などを開発してまいりました。

今後 NTT では、新開発の「小型インテリジェントマイク」によって、従来技術では困難であった音声サービスを実現し、様々な端末と音声を利用したサービスの飛躍的な発展に貢献してまいります。(図 2)

※「モバイル音声認識向け背景雑音抑圧技術*2」、「音声会議・録音機器向け背景雑音抑圧技術*3」

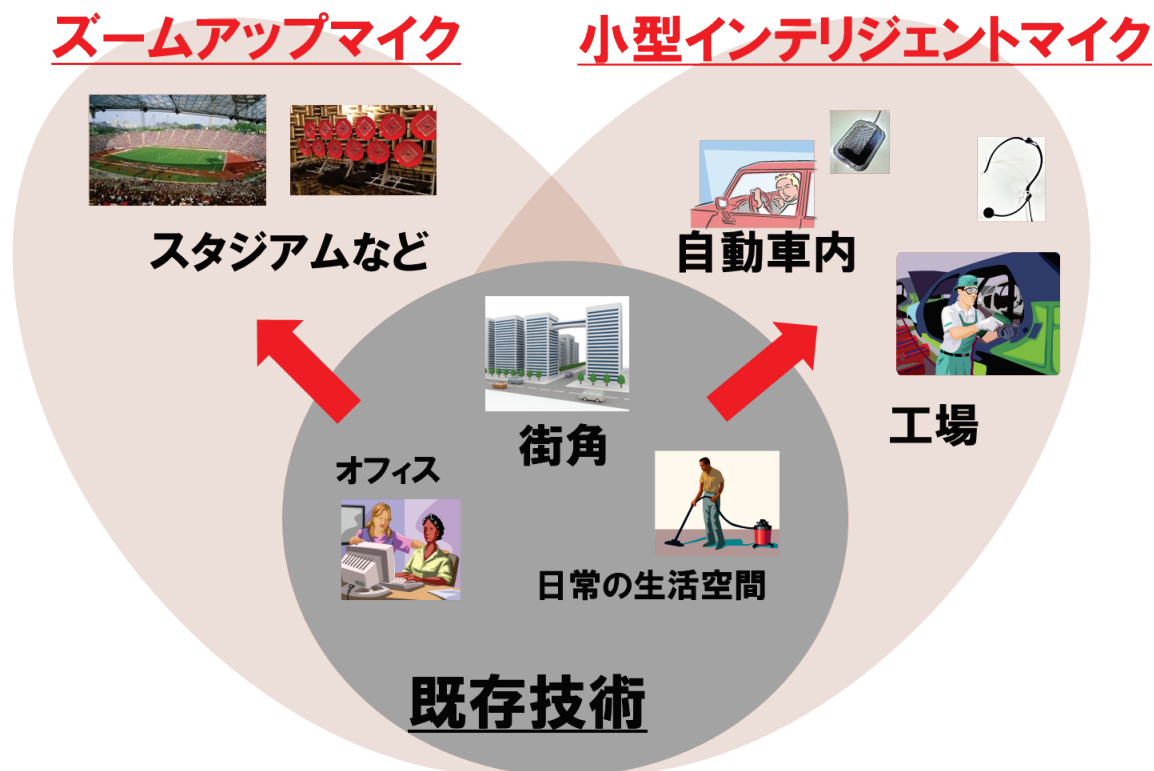
【販売について】

エヌ・ティ・ティ・アドバンステクノロジー株式会社(以下:NTT-AT、本社:神奈川県川崎市、代表取締役社長:花澤隆)は、今回 NTT が開発した「小型インテリジェントマイク技術」をもとにした、ソフトウェアライブラリ「インテリジェントマイクライブラリ(仮称)」の販売を 2014 年 10 月より開始します。また、その他、音響処理技術も取得し、本ライブラリを、これら各種技術と組み合わせ、様々な利用シーン・ハードウェア構成における音声を利用した商品／サービス／ソリューションとして提供していく予定です。

(図1)＜利用イメージ＞



(図2) <NTT の技術が可能とする音響処理の適用範囲>



- *1 凹型反射板の前に準最適に配置された多数のマイク素子と、それぞれのマイク間に生じる位相／振幅の差を利用することで、任意の位置の目的音を強調して集音する技術(2014 年 4 月 16 日ニュースリリース) <http://www.ntt.co.jp/news2014/1404/140416a.html>
- *2 音声らしさを表す統計音声モデルとユーザーの発話の経過に応じて変化する特徴量を用い、雑音下での音声認識精度を大きく改善する技術
- *3 通話向け用途に、比較的長時間(数秒～10 秒程度)の雑音データからその雑音のスペクトルの時間変化推定することで、雑音の時間変化にも対応した雑音除去技術

<本件に関するお問い合わせ先>

日本電信電話株式会社
サービスイノベーション総合研究所 企画部広報担当
TEL: 046-859-2032
E-mail: randd@lab.ntt.co.jp

エヌ・ティ・ティ・アドバンステクノロジー株式会社
商品お問合わせセンタ
TEL: 0120-057-601
E-mail: sales@ml.ntt-at.co.jp

【NTT】インテリジェントマイク技術の開発について

1. 開発の背景

これまで、工場内や工事現場等の機械騒音下や高速道路で走行中の自動車内など高騒音下において、音声認識やハンズフリー通話を行った場合、音声認識率の著しい低下や通話音声の劣化が生じ、実用が困難でした。

2. NTT が開発した技術のポイント(3点)

「小型インテリジェントマイク」は、音の到来方向、周波数特性、時間的な変動特性を最大限に活用し、目的とする音声とそれ以外の雑音を高精度に推定して分離することを可能とするものです。

これにより、従来の技術では実現が困難であった“小型”かつ“2～3 個のマイクロホン素子”でも、目的音声をほとんど劣化させることなく、周囲雑音のパワーを約 1/10000 まで低減させることが可能となり、100dB といった超高騒音下において、高精度な音声認識や高品質な通話を実現します。

<ポイント1> ビームフォーミングとスペクトルフィルタ処理のハイブリッド構成

目的音声の到来方向に指向性向け、その音を強調するビームフォーミング処理に加え、雑音成分を推定して、周波数スペクトル領域で抑圧するスペクトルフィルタ処理を組み合わせる構成としました。

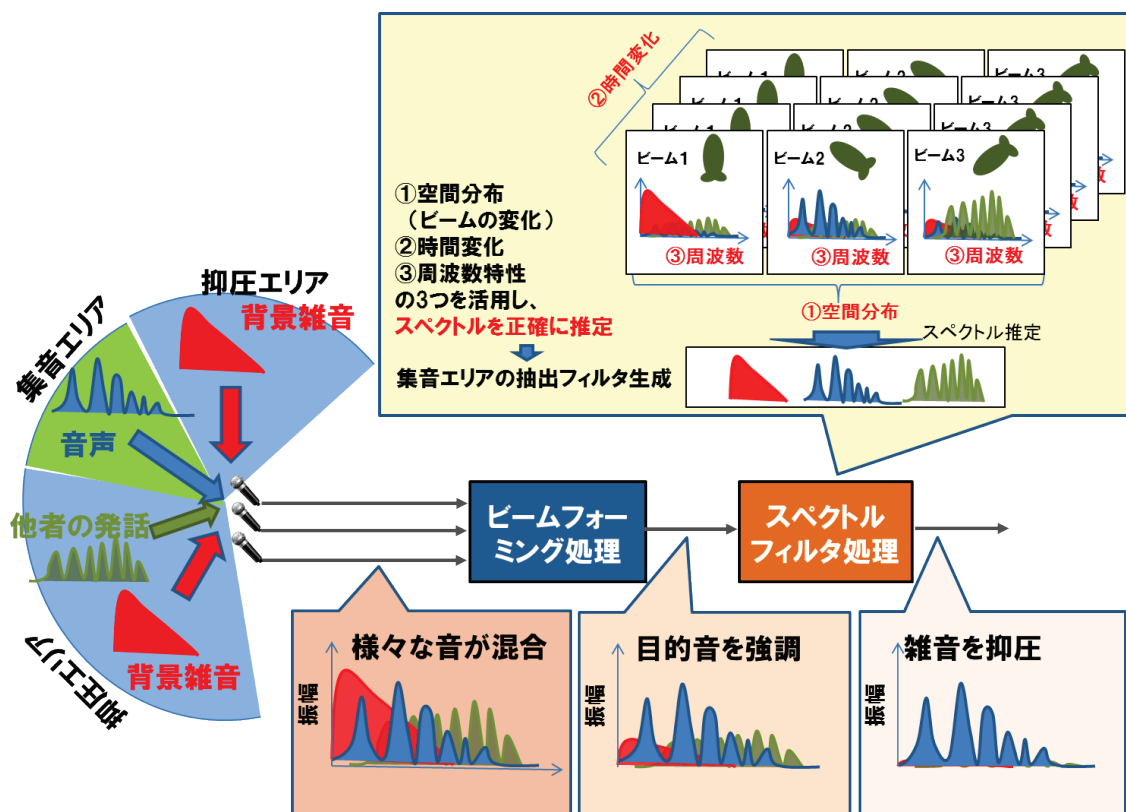
少数のマイクで形成するビームフォーミングでは鋭い指向性を形成することができず、十分に目的音声を強調することができませんが、後段のスペクトルフィルタ処理で補うことにより、多数のマイクを利用するビームフォーミングに匹敵する性能が得られます。

<ポイント2> 雑音の空間分布推定

ビームフォーミングは、目的とする音声を狙うだけでなく、レーダーのようにそれ以外のあらゆる方向に指向性を形成し、方向ごとの雑音の空間分布を推定します。目的音と雑音の空間分布が異なっていれば、雑音成分を従来よりも正確に推定することができます。

<ポイント3> 雑音の時間変動特性を利用

雑音には、エアコンの音のように定常的な雑音、BGM や他者の話し声のように時間変動する雑音があります。推定した雑音の空間分布と時間変動特性に応じた方法で、雑音の周波数スペクトルを高精度に推定します。



【参考】これまでの雑音抑圧技術

雑音抑圧技術には、単一のマイクロホンを用いる技術と、複数のマイクロホンを用いる技術があります。
以下の課題をクリアしたものが、今回開発した「小型インテリジェントマイク」です。

＜課題：単一のマイクロホンを用いる技術＞

音声発話のない区間の信号から背景雑音の周波数特性を推定し、入力信号から推定した雑音成分を差し引くことで目的とする音声を強調することができます。

オフィス環境やエアコンの騒音など、騒音レベルがそれほど高くなく、時間変動の少ない背景雑音環境での利用が可能です。しかし、工場内の機械騒音のように雑音が時間的に大きく変動したり、高速走行中の自動車内のように定常でも騒音レベルの高い場合に適用すると、処理音声に大きな劣化が生じてしまう問題がありました。

＜課題：複数のマイクロホンを用いる技術＞

複数マイク用いるビームフォーミングでは、音源位置の空間的違いを利用するために、雑音の性質によらず抑圧できるメリットがありますが、高い雑音抑圧量を得るには、数 10 個のマイクロホンを使う大型の装置が必要で、小型の装置で高い抑圧量を得ようとすると、目的音まで劣化してしまう課題がありました。

		非定常な雑音の抑圧 高騒音下での音質	雑音抑圧量	コストやサイズ
従来	単一マイクロホンの雑音抑圧 (スペクトルサブトラクション)	×	○	◎
	少数マイクロホンの雑音抑圧 (ビームフォーミング)	○	×	○
	多数マイクロホンの雑音抑圧 (ビームフォーミング)	○	◎	×
今回	小型インテリジェントマイク	○	◎	○

【NTT-AT】ソフトウェアライブラリ「インテリジェントマイクライブラリ(仮称)」の販売開始について

NTT-AT はこのたび、「小型インテリジェントマイク技術」を NTT より取得し、2014 年 10 月(予定)より、ソフトウェアライブラリ「インテリジェントマイクライブラリ(仮称)」を販売いたします。お客様のご利用形態や開発製品に合わせたカスタマイズサービスも併せて提供いたします。さらに、同ライブラリを利用したサービス・商品をリリースすると共に、お客様のご要望に応じて、各種先端技術と組み合わせたソリューションを提供していく予定です。

今後、NTT-AT では、NTT グループ企業や関連企業との連携により、自動車、工場、工事現場、航空機等での通話、音声認識用途での営業展開を図ってまいります。

＜販売に関する概要＞

- ・販売開始:2014 年 10 月予定
- ・販売形態:ソフトウェアライブラリおよびカスタマイズ
- ・販売方法:NTT-AT が直接販売いたします。
- ・販売価格:お客様のご利用形態・開発商品等により細かく対応いたしますので、NTT-AT にお問い合わせください。
- ・販売エリア:日本国内から提供開始
- ・販売目標:100 万／年ライセンス程度の販売を計画しています。

＜利用シーンの拡大に向けたラインナップ＞

NTT-AT では、お客様のご利用環境、用途によって、単一マイクでのご利用を希望される場合には、NTT から技術を取得した、単一マイク向けの「背景雑音抑圧技術」のライブラリも提供可能です。

また、固定小数点 DSP 向けライブラリ、組み込み CPU 向けライブラリ、PC 向けライブラリのラインナップがある他、お客様のご利用環境に合わせたポーティングや処理のカスタマイズにも対応いたします。さらに、エコーキャンセラ、自動音量調整、コーデック、VoIP 関連技術等、音響関連のソフトウェアライブラリの提供を総合的に行います。