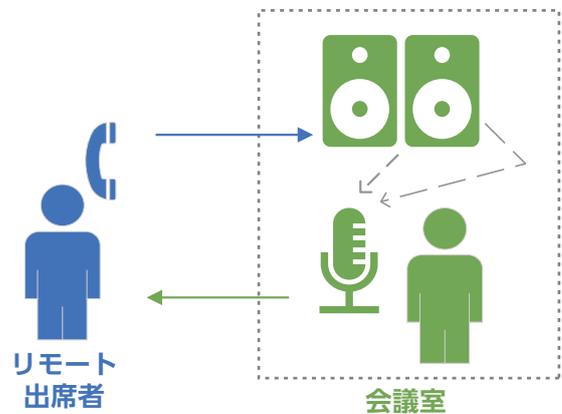


右図のようなテレビ会議やビデオ会議の典型的な問題として、リモートの参加者に自分の発言の遅延したエコー(反響)が聞こえてしまうというものが挙げられます。これは、ファーエンド (Far-End: 遠端) 信号とも呼ばれるリモート話者の音声、その部屋のスピーカー、部屋のインパルスレスポンス、マイクを介して再生されるために起きる現象です。

AEC(アコースティック・エコー・キャンセラー)は、ニアエンド (Near-End: 近端) と呼ばれるローカルマイクのシグナルからシミュレートされたエコーを差し引くことで、これを回避することを目的としています。

AECの性能と実装は様々です。本書では、Allen & Heath AHMオーディオ・マトリクス・プロセッサで使用するために設計されたシステムについて説明します。機能要素について説明し、その後、ユーザーインターフェースの概要を説明します。



M-AHM-64 拡張モジュール

このプロセッシング拡張モジュールは、12chのAECを搭載しており、各チャンネルは150msのエコーキャンセルが可能です。このモジュールは、デュアルコア ARM Cortex A9 プロセッサをそれぞれ搭載した2つのXilinx FPGA デバイスを利用しています。

以下の図は、システムの主要コンポーネントを示しています。

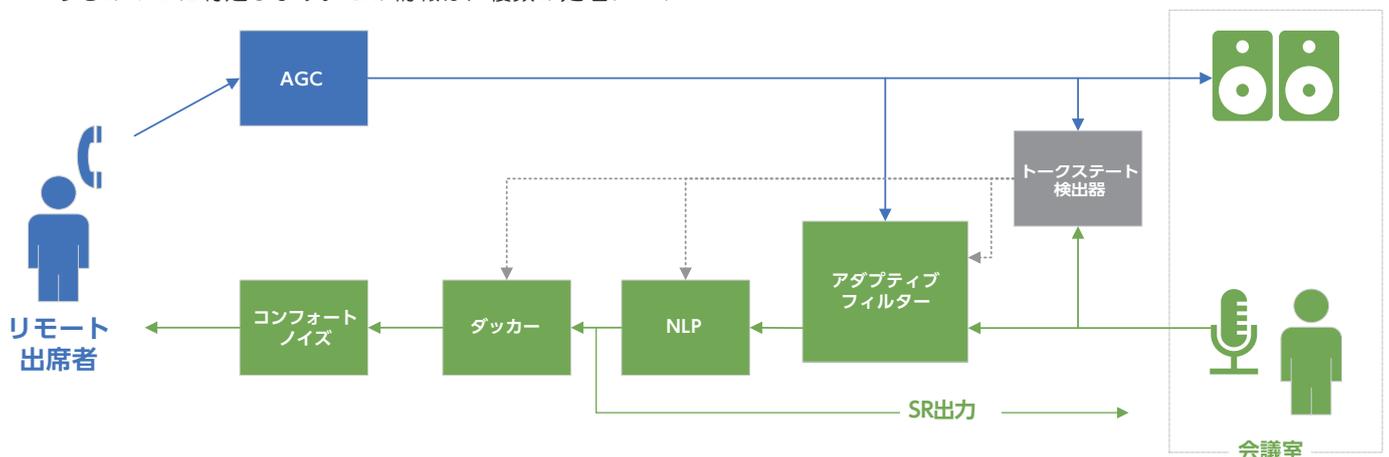
アダプティブ・フィルタ (Adaptive Filter) は、常に更新され、部屋のインパルスレスポンスに一致させ、エコーをシミュレートします。

このプロセスを補助し、総合的に結果を改善する他のブロックは以下の通りです。

- ・ **AGC (自動ゲイン制御)** は、リモート側からの通話のレベルを一定に保つようにファーエンドのゲインを調整します。信号に音声が含まれていない限り、ゲインは変更されません。
- ・ **トークステート検出器** は、音声の有無とダブルトーク (ファーエンドとニアエンドの両方が話している状態) を識別します。ニアエンドが話している、ファーエンドが話している、両方が話している、どちらも話していない、の4つの会話状態のうちの1つに特定します。この情報は、複数の処理ブロッ

クに渡されます。例えば、ダブルトークが検出された場合、ファーエンドのエコーとニアエンドの話者を区別できないため、アダプティブ・フィルタは更新されません。

- ・ **NLP (ノンリニアプロセッサ)** は、アダプティブ・フィルタでキャンセルできなかった残留エコーを減衰させる追加の処理を行い、アダプティブ・フィルタが収束する間、主なエコーを低減します。また、バックグラウンドノイズ低減機能も備えています。
- ・ オプションの **ダッカー** は、ファーエンドのみが通話中の場合、通話状態の情報を使用して残留エコーのレベルを低減します。
- ・ **コンフォートノイズ (Comfort Noise)** は、リモート側の通話者が回線が切断されたと誤認する可能性のある完全な無音状態を回避するために、リターンパスにノイズを追加します。ノイズのスペクトルは、自然で不快でない音になるように調整されています。



AEC システムの主なコンポーネント

セットアップとルーティング

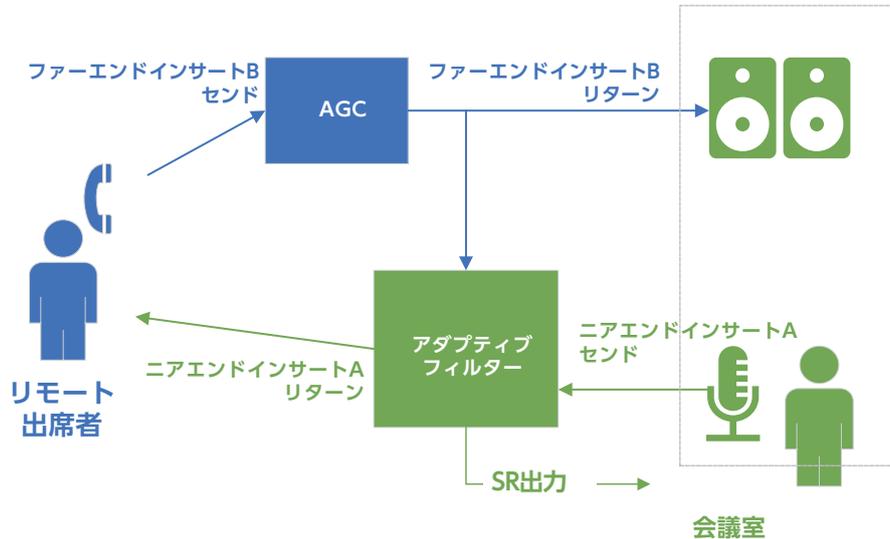
AEC はチャンネルインサートを使用して割り当てられます。一方のインサートはファーエンドの入力チャンネルに、もう一方はニアエンドの入力チャンネルに配置されます。ファーエンドのインサートは Insert B に限定されており、ダイナミクス処理の後になるように設定されています。ファーエンドのソースとスピーカーの間にノンリニア処理が存在する場合、AEC は機能しません。

同様の理由により、ニアエンドのインサートは Insert A に限定されており、ダイナミクス処理の前になるように設定されていま

す。ニアエンドの入力チャンネルのゲートは無効にする必要があります。

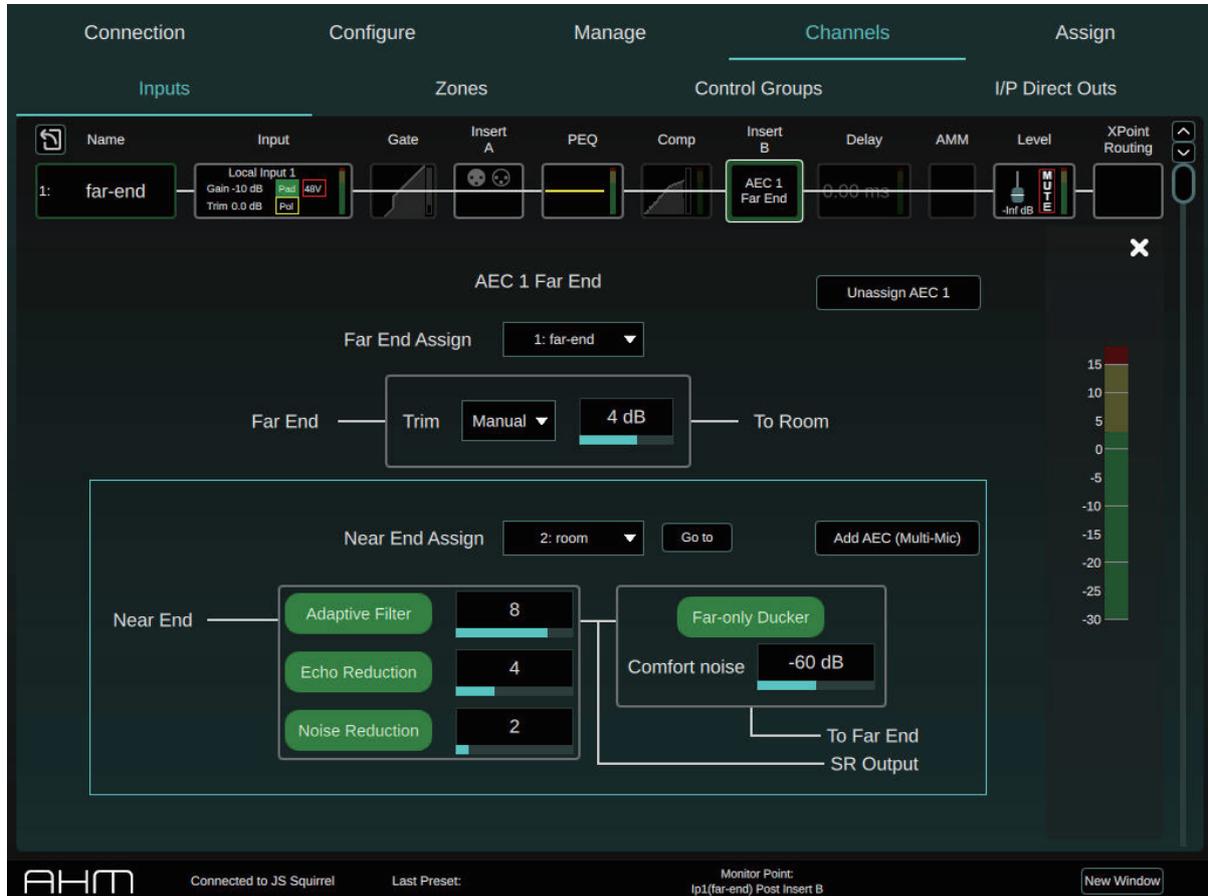
正しいインサートのルーティングは、以下の略図に示されています。

SR Output (サウンド・レインフォースメント出力) も用意されています。通話状態に応じたゲートやコンフォートノイズの処理前のニアエンド出力を追加で供給するもので、室内の音声拡声に利用できません。



AHM システムにおける AEC ルーティング

ユーザーインターフェース



ファーエンド側のインサートから始めます。ファーエンド側のインプットチャンネルの Insert B に進み、**Assign AEC (far-end)** をクリックし、挿入する 12 の AEC プロセッサから 1 つを選択します。すでに使用中の AEC はグレー表示になります。AEC の設定とパラメーターが表示されます。

！ AHM System Manager とファームウェア V1.10 以上が必要で

ニアエンド側を割り当てるには、上記の画面の **Near End Assign** ドロップダウンでソースを選択します。または、ニアエンド側の入力チャンネルで Insert A を選択し、**Assign AEC (near-end)** を選択します。

！ ファーエンド側のインサートからニアエンド側のスピーカーへの経路には、**ダイナミックプロセッシング** (コンプレッサーまたはゲート) を配置しないでください。

下部パネルの **Go to** ボタンをクリックして、ニアエンド側とファーエンド側のインサート画面を切り替えます。ファーエンドの画面には AEC リファレンスマーターが表示され、ニアエンドの画面には AEC 出力メーター (エコーキャンセルの適用後) が表示されます。

！ ニアエンドインサートで検出されるエコーのレベルがファーエンドインサートのリターンよりも高い場合、AEC は機能しません。適切なゲイン構造と適切なローカルマイクゲインが設定されていれば、これは問題にはなりません。

複数マイクシステムを設定するには、ニアエンドインサート画面で **Add AEC (Multi-Mic)** をクリックします。リンクされたすべての AEC プロセッサを選択し、ニアエンド・インプットチャンネルを適切に割り当てることができます。マイクミックス (または会議システム出力) に単一の AEC を使用するのではなく、各マイクソースに AEC プロセッサを使用することで、全体的なパフォーマンスが最適化されます。

室内での音声増幅を目的とする場合は、処理済みのニアエンド・インプットチャンネルではなく、SR Output を使用することをお勧めします。Zone の **XPoint Routing** 画面に移動し、適切な **AEC SR リターン** チャンネルのレベルを調整します。この方法では、エコーキャンセリング性能と音声増幅のバランスが最も良くなります。SR Output では、**アダプティブ・フィルタ** によるレイテンシーが発生しますが、コンフォートノイズやゲートは適用されません。

別の方法として、マイクの音源を 2 つ目の入力チャンネルに分割して音声増幅に使用する方法もあります。これにより、AEC プロセッサのレイテンシーやその他の影響を受けずに、最高の音声再生品質が得られますが、エコーキャンセリングのパフォーマンスに悪影響を及ぼす可能性があります。

コントロールとパラメータ

Trim を調整して、メーターが典型的なファーエンド・ソースで 0dB 付近を示すようにします。低レベルのファーエンド信号では、AEC のパフォーマンスが低下します。

Auto を選択すると、AGC が有効になり、手動トリムが無効になります。

Adaptive Filter を選択すると、**アダプティブ・フィルタ** が有効になり、その速度を制御できます。許容できるパフォーマンスが得られる最低の値に設定します。

Echo Reduction を選択すると、エコーをさらに低減する NLP ブロックの一部が有効になり、制御できます。許容できるパフォーマンスが得られる最低の値に設定します。高い値を設定すると、ファーエンドの受信機で歪みが生じる場合があります。

Noise Reduction を選択すると、会議室内の定常的なバックグラウンドノイズを低減する NLP ブロックの一部が有効になり、制御できます。ここでも、有用なバックグラウンドノイズ低減が

得られる最低値を設定します。

特に自然界に散発的に発生する低周波ノイズの除去を改善するには、ニアエンド・インプットチャンネルの HPF を使用します。

Far-only Ducker は、トークステート検出器によって駆動する減衰ステージを可能にします。ファーエンドのみが通話している場合、ニアエンドの信号は減衰されます。

Comfort Noise は、ファーエンドの通話者に送信される音声ノイズのレベルを制御します。NLP ブロックの有効性を向上させるため、ノイズを完全にオフにすることはできません。

技術解説 - トークステート検出器

大抵の AEC システムブロック図には、少なくともダブルトーク検出器 (DTD) が記載されており、多くの場合、1 つ以上の音声活動検出器 (VAD) も記載されています。これらはトークステート検出器 (TSD) と呼ばれる 1 つのブロックに統合されています。DTD、ひいては TSD の主な機能は、ニアエンドマイクで拾った信号にニアエンドの音声とファーエンド信号のエコーの両方が含まれているかどうか、つまりダブルトークかどうかを判断することです。ニアエンド側の室内で音声が発せられているときにアダプティブ・フィルタが作動すると、アダプティブ・フィルタは発散してしまうので、これは可能な限り回避しなければなりません。

AHM の仕様では、バンドパス・フィルタを通したニアエンドおよびファーエンドの信号の統計的性質を推定し、それを基に TSD の判断を行います。

技術解説 - NLP

アダプティブ・フィルタは室内インパルス応答の識別に有効ですが、若干の残響エコーが残ります。これは、室内の反響が変化した場合や、システムが起動したばかりの場合に特に顕著です。このため、残留エコー抑制システムを使用して、このような状況を補完し、全体的なパフォーマンスを向上させます。

このシステムでは、スペクトラル減算法が使用され、ファーエンドとアダプティブ・フィルタの誤差信号の両方が周波数領域に変換されます。これは既存の技術であり、主な課題は、このプロセスでしばしば発生するアーティファクトや歪みを最小限に抑えること、および差し引く値の正確な推定値を算出することです。

バックグラウンドノイズの低減

バックグラウンドノイズは明瞭度を低下させ、残差予測システムを妨害する可能性があります。バックグラウンドノイズのスペクトルは、パワースペクトル最小値を追跡することで推定できます。ニアエンドの音声バックグラウンドノイズの推定に影響を及ぼさないように、ニアエンドマイクがアクティブではないと判断された場合 (すなわち、トークステート検出器からの短期 SNR 推定値が低い場合) のみ、推定値が更新されます。

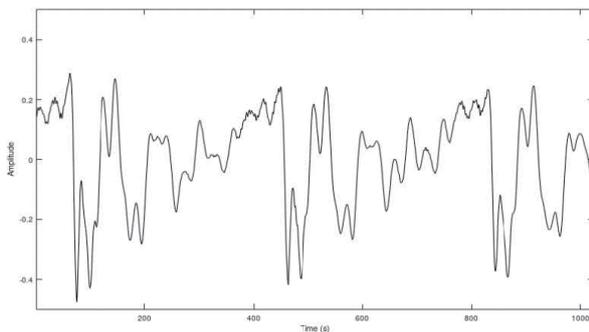
技術解説 - アダプティブ・フィルター

周波数領域の実装

アダプティブ・フィルターは FIR (有限インパルス応答) フィルターであり、システムのインパルス応答が長くなるにつれ、フィルターの長さもそれに合わせて長くする必要があります。一般的な AEC のアダプティブ・フィルターの長さが 100 ミリ秒を超える場合、このような畳み込みフィルターを実装するには、非常に大きな演算能力が必要となります。演算要件を削減する技術として、フィルターを周波数領域で適用する方法があります。時間領域での畳み込みは周波数領域での乗算と等価であるため、この方法は有益です。このような場合、フィルターを適用する処理の削減により、正フーリエ変換および逆フーリエ変換の負担はすぐに克服できます。この手法のデメリットは、有効な長さの変換を行うと自動的に遅延が発生することであり、これは純粋な時間領域解では発生しないものです。フィルターの分割技術により、この遅延を許容可能なレベルまで低減することができ、なおかつ計算上の利点のほとんどを維持することができます。

NLMS 適応

ノーマライズ適応は、信号のエネルギー量に基づいて、周波数ごとにフィルターの更新速度を動的に調整します。ノーマライズ化を行わない場合、システムの動作レベルに非常に敏感になります。小さな信号ではフィルターに有意義な変化をもたらすことができず、大きな信号ではフィルターが不安定になる可能性があります。



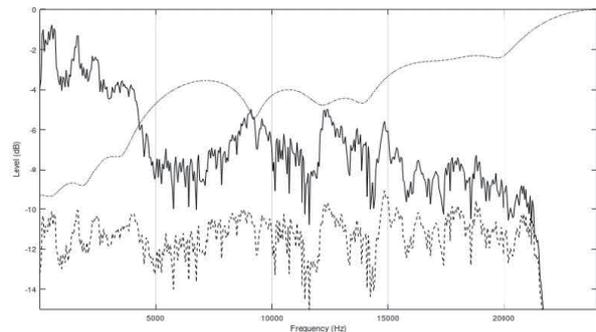
信号のホワイトニング

ホワイトニングはアダプティブフィルタリングの確立された技術であり、入力信号にプリフィルタリングを適用して、より平坦なスペクトルにします。これは、信号がホワイトノイズである場合にアダプティブ・フィルターが最も速く収束するのと同じ理由で有利です。信号がよりホワイトノイズに近ければ近いほど、真のエコー信号以外のものとの誤った相関が発生する可能性が低くなり、フィルター更新プロセスにおける相関は、フィルター係数を収束に向かわせる可能性はるかに高くなります。

ホワイトニングされた信号に基づいて適応する必要があるため、このアダプティブ・フィルターからの出力、つまり「誤差」信号は、必要なエコーキャンセラー処理された音声ではありません。そのため、ホワイトニングされた経路から生成された係数は、エコー推定値を生成し、実際の AEC 出力を生成するために使用される 2 番目のフィルターにコピーする必要があります。

以下のグラフは、音声サンプル例と対応する周波数スペクトル、また結果として得られた AR モデルによるホワイトニング・フィルターの周波数特性を示しています。

ホワイトニングの結果から、このプロセスがスペクトルを平坦化することに有効であり、アダプティブ・フィルターにとってより理想的な入力信号を提供することは明らかです。またホワイトニングされたスペクトルは明らかにパワーが減少しますが、NLMS 正規化により、システムには影響を与えません。



左：音声サンプル。右：音声サンプルのスペクトル（実線）、ホワイトニング・フィルターの応答（点線）、ホワイトニングされたスペクトル（破線）。