



L-ISA ハイパーリアルサウンド：フィルシステムにおけるカバレッジとスペイシャルレンダリングの最適化

L-ISA テクノロジーは、共有カバレッジが限られているラウドスピーカーシステム（フロントフィルやアンダーバルコニーフィルなどの分散型フィルシステム）に対して、「スペイシャル（空間的な広がりのある）フィル」を作成するオプションを提供します。このアプローチでは、まずシーンシステムのラウドスピーカーの仮想的な複製を作成してクロスカバレッジを復元し、次にゲインベースのアルゴリズムを使用してオーディオオブジェクトを配置します。これにより、カバレッジとレベルの一貫性を確保しながら、オーディオオブジェクトの分離と視聴覚の一貫性を改善します。

L-ISA ハイパーリアルサウンド：カバレッジとパンニング

L-ISA テクノロジーの基本原則の一つは、観客カバレッジ*をスペイシャルレンダリングのステップから分離することです。カバレッジは主に設計によって保証され、同じ観客部分に対応するフルレンジソースと、観客エリア全域にわたるカバレッジを保証するために設けたフィルシステムによって定義されます。そうすることで、選択したパンニングアルゴリズムやオブジェクトの位置に依存せずにカバレッジが保証されます。しかし、これまでフィルシステムでしかカバーされていないエリアでは、空間的な広がりのある音の再現性は得られませんでした。



L-ISA テクノロジーは、これまでモノラル信号を供給していたフロントフィルやアンダーバルコニーフィルなどの分散フィルシステムでカバーする領域に、空間的な広がりのあるサウンドを作り出す新たなスペイシャルフィルソリューションを提供します。このオプションでは対応する観客エリアにいるリスナーが少なくとも 3 台のラウドスピーカーのカバレッジエリア内にあることが必要です。この要件に基づき、最初のステップとしてクロスカバレッジを復元するためのバーチャルラウドスピーカーを作成します。次にこのバーチャルラウドスピーカーにスペイシャルレンダリングアルゴリズムからの信号を供給します。

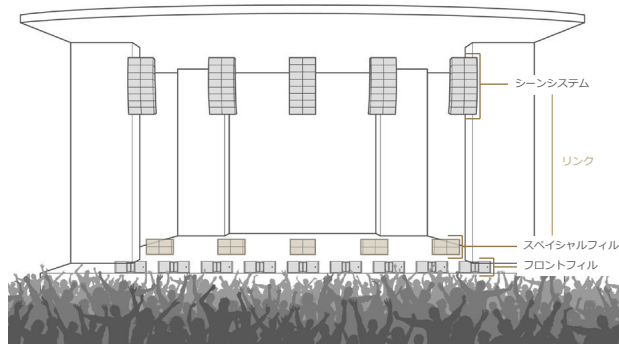
L-ISA テクノロジーのスペイシャルフィルソリューションは、大規模イマーシブシステムに必要な不可欠な 3 つの特徴と利点である、カバレッジ、空間的なアンマスキング*によるオブジェクトの分離、視聴覚の一貫性*の良好なバランスを生み出します。

カバレッジ、空間的なアンマスキング、視聴覚の一貫性のバランス

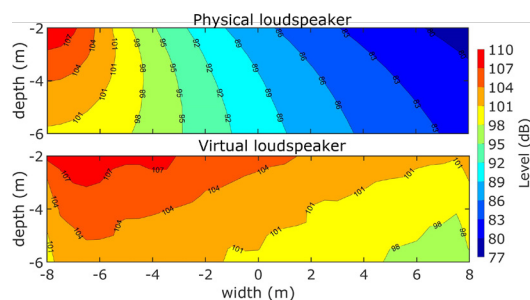
ラウドスピーカーシステムの仮想化

スペイシャルフィルの目的はシーンシステムフルレンジソースのカバレッジを向上させることであり、分散フィルシステムのソースを組み合わせることでシーンシステムの仮想レプリカを作り出すことでこれを実現します。仮想ラウドスピーカーは、以下のパラメータを用いてゲイン分布を最適化した遅延ベースのアルゴリズムを用いて作成されます。

- 仮想距離（単位：m）：フィルシステムから仮想ラウドスピーカーまでの距離
- ゲイングラデーション（単位：dB）：作成されたすべての仮想スピーカーのゲインの高低差



右図は「フロントフィルシステム・アウターハウスレフト・フィジカルラウドスピーカー」の SPL マッピング (フロントフィルのカバレッジエリア) と「ハウスレフトシーンの仮想レプリカラウドスピーカー」の SPL マッピングを比較しています。客席反対側までの間の SPL 差は、単一のラウドスピーカーと比べてスペイシャルフィルでは大幅に減少し、スペイシャルフィルのカバレッジエリアはフロントフィルのカバレッジエリア全体に拡大します。シミュレーションの詳細は付録 1 を参照してください。

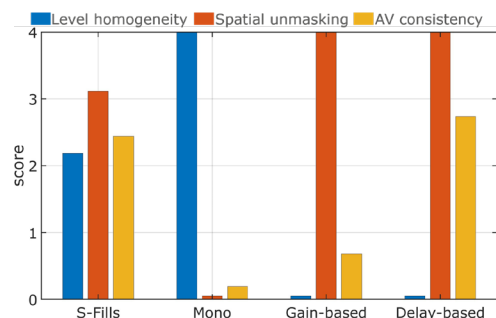


フロントフィルに対するメリット

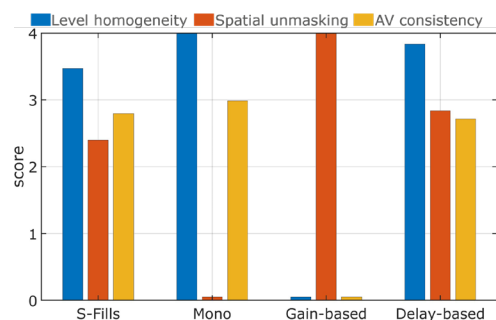
右図はフロントフィル用の様々なスペイシャルレンダリングソリューションを比較したものです。棒グラフはパフォーマンス指標を表します。(詳細は付録 2、3、4 を参照してください。)

この比較は、L-ISA (S-Fills) のスペイシャルフィルソリューションのみが、3 つの基準 (レベルの均質性、空間的なアンマスキング、視聴覚の一貫性) のバランスが取れていることを示しています。

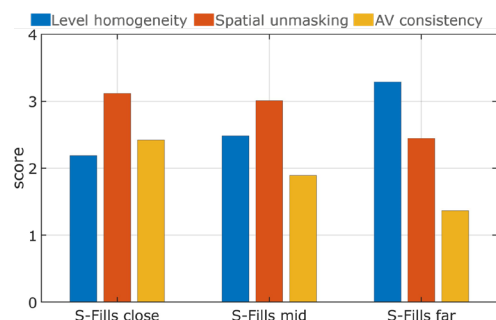
「Mono」ソリューションは、レベルの均質性は最も優れていますが、空間的なアンマスキングは存在せず、視聴覚の一貫性も限られています。「Gain-based」のソリューションでは、レベルの均質性と視聴覚の一貫性を犠牲にして、空間的なアンマスキングを改善しています。これは、フロントフィルのラウドスピーカーとして使用するフィジカルフルレンジソースのカバレッジが共有されていないためです。



「Delay-based」のソリューション (Wave Field Synthesis) は、空間的なアンマスキングと視聴覚の一貫性には最も適していますが、レベルの均質性を確保することができません。舞台手前のオブジェクトポジションでは、ディレイベースのソリューションはすべてのエネルギーを少数のラウドスピーカーに集中させる傾向があり、結果的にレベルの均質性に問題が生じます。



2 番目の図は、3 つの基準に対するスペイシャルフィルソリューションのパラメータの影響を示します。S-Fills close の設定 (仮想距離 5m、ゲイングラデーション 8dB) では、カバレッジ、空間的なアンマスキング、視聴覚の一貫性のバランスが最も良くなっています。仮想距離が長くなり (S-Fills mid は 8m、S-Fills far は m16)、ゲイングラデーションが小さくなると (S-Fills mid は 7dB、S-Fills far は dB5)、カバレッジは向上しますが、まず視聴覚の一貫性が損なわれ、次に空間的なアンマスキングが損なわれます。視聴覚の一貫性については、付録 5 を参照してください。



アンダーバルコニーに対するメリット

同様の分析をアンダーバルコニーフィルラウドスピーカーシステムについても行いました。(テストシナリオについては付録 1 を参照してください) 「Spatial-fills」と「Delay-based」は「Mono」と「Gain-based」を上回る結果となりました。「Mono」はカバレッジエリア (ステージから 16m、深さ 8m のエリア) において、良好なレベルの均質性と視聴覚の一貫性を確保しています。「Gain-based」は空間的なアンマスキングには優れていますが、レベルの均質性と視聴覚の一貫性は損なっています。

用語集

オーディオオブジェクト : オーディオ入力とその空間的な位置関係などのプロパティを記述したメタデータを関連付けたもの。

カバレッジ : ラウドスピーカーシステムが許容可能な周波数応答のバリエーションで直接音を提供する範囲のこと。

視聴覚の一貫性 : オブジェクトやフォーマーに関連する聴覚と視覚の手がかりを、同じ場所から発信されたものとして認識する能力。

空間的なアンマスキング : 複数の音のそれぞれの空間的な起源を正確に認識することで、お互いをマスクさせない能力。

付録 1 : テストシナリオ

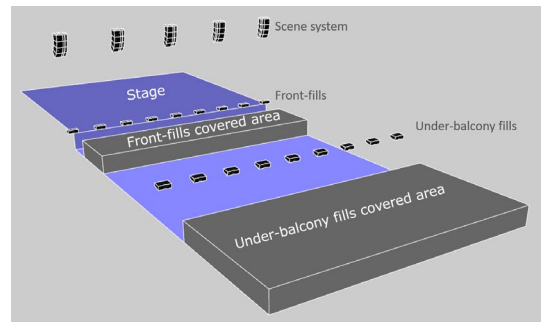
テストシナリオでは、16 x 16m のステージを備えた、幅 16m の典型的なシューボックスの会場を想定しています。シーンシステムはステージ横幅方向にまたがる 5 つのフルレンジソースで構成しています。これを補完するものとして、一定の間隔 (2m) で配置した 9 x Kara II ラウドスピーカーで構成するフロントフィルシステムと、ステージから 16m 離れた場所に同様のアンダーバルコニーシステムを配し、それぞれが特定のエリアをカバーしています。

ステージ横幅と奥行にわたりグリッドを作成し、テストオーディオオブジェクトを配します。

- 横幅 : 0 (中央)、2、4、6、8 m
- 奥行 (舞台奥へ) : 2.5、4、5.5、8、12、16 m

L-ISA スペシャルフィルソリューションは以下のチューニング値を使用 :

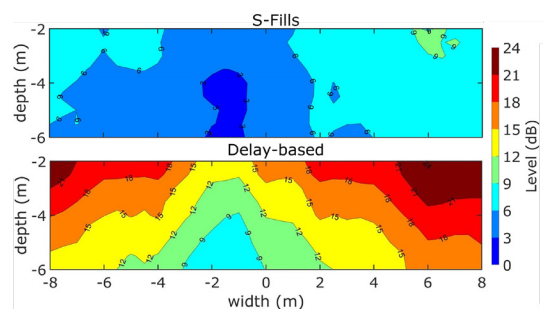
- フロントフィル : 仮想距離 5 m (ステージ横幅の 1/3)、ゲイングラディエーション 8 dB
- アンダーバルコニーフィル : 仮想距離 16 m (シーンからアンダーフィルまでの距離)、ゲイングラディエーション 4 dB
- さらに多くの構成をテストしましたが同様の傾向と結果が得られたためここでは説明しません。



付録 2 : レベルの均質性の評価

レベルの均質性は、Soundvision 伝搬モデルを用いて、1 kHz から 10 kHz の帯域で生成された SPL から得られる基準です。各観客ポジションとスペーシャルレンダリングソリューションについて、レベルの均一性は、レンダリングされたすべてのオーディオオブジェクトポジションの中で最も高い SPL と最も低い SPL の差として計算されます。この値が小さいほど、カバーエリア内におけるミックスの一貫性を保てます。

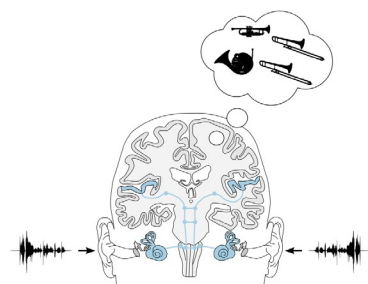
最初の図はテストしたフロントフィル構成におけるスペーシャルフィルのレベル均質性のマッピングを示します。(付録 1 を参照) 2 番目の図は遅延ベースアルゴリズムのレベル均質性を表しており、テストしたオーディオオブジェクトポジション全体で高いレベルの差 (最大 24 dB) を示しています。この場合、舞台手前のオブジェクトの位置は、オフアクシスラウドスピーカーの低いゲイン値に対応し、個々のラウドスピーカーのカバレッジが限られているため、低いレンダリング SPL になります。



付録 3 : 空間的なアンマスキングと視聴覚の一貫性の評価

空間的なアンマスキングと視聴覚の一貫性を、その精度が実証されている聴覚モデルを用いて評価します。聴覚モデルにはコンサート中にリスナーの両耳に到達する音波に対応するバイノーラル信号を入力します。無響室で測定した頭部伝達関数は、環境の残響を考慮せず、直接音のみに集中した自由音場の状況をシミュレートするために使用されます。聴覚モデリングツールボックスの 2 つのモデルを使用しています :

- wierstorf2013 は、視聴覚の一貫性の評価のために、オーディオオブジェクトの定位誤差の推定値を提供。
- jelfs2011 は、空間的なアンマスキングのために、ターゲットとインターフィアラー (妨害し合う) が空間的に分離したときのターゲットの音声明瞭度の増加の推定値を提供。



ベルリン工科大学による Kemar 測定、距離 3 m、[ここから入手できます](#)。

付録 4 : 品質スコア

同じ尺度や単位を持たない基準に対する異なるソリューション間の比較を容易にするために、各基準に対して品質スコアシステムを提案し用いています。

レベルの均質性は、すべての観客ポジション間の SPL 差（最悪のシナリオ）の 95 パーセンタイル（単位 dB）として計算。オブジェクトが聞こえないポジションがあってはならないため、最悪のシナリオを考慮する必要があります。

空間的なアンマスキングは、すべてのオブジェクトペア（ターゲットセンター、同じ奥行で左右の位置が異なるインターフィアラー）とすべての観客ポジションの中央値（単位：dB）。

定位誤差は、全てのオブジェクトポジション（ステージ上の幅と奥行）と全てのオーディエンスポジションの中央値（単位：°）。定位誤差が小さいほど、視聴覚の一貫性が高いことになります。

品質スコアは以下のしきい値を用いて定めます。（品質スコアが高いほど良い）

品質スコア	4*	3*	2*	1*
レベルの均質性 (単位：dB)	≤6	>6	>9	>12
空間的なアンマスキング (単位：dB)	>4	>3	>2	>1
定位誤差 (単位：°)	<5	<10	<15	<20

付録 5 : 仮想距離の調整とオーディオ-ビジュアル品質のトレードオフ

この付録では、フロントフィルシステムのカバレッジエリア内でリスナーが経験する定位誤差が、スペイシャルレンダリングソリューションとステージ上のオーディオオブジェクトの奥行によってどのように変わるかを調べます。ステージ上のオブジェクトは、定位誤差が最も大きいハウスレフトの外側に配置しています。最初の図は、スペイシャルレンダリングソリューション間の定位誤差の分布（菱形：中央値、縦棒：25 パーセンタイル、75 パーセンタイル）を比較したものです。

「Mono」と「Gain-based」のソリューションは、ステージ上のオーディオオブジェクトの奥行が浅いときと深いときに、それぞれ非常に大きな定位誤差を示します。「Spatial-fills」と「Delay-based」ソリューションはどちらも似たような性能で、ステージ上のオブジェクトの深さにはほとんど依存しません。中央値の定位誤差は目標の閾値である 7.5 度を超えています。2 つ目の図は、「Spatial-fills」ソリューションの仮想距離パラメータ（ステージ前方からの仮想ラウドスピーカーまでの距離）について複数の値を比較したものです。

- S-Fills close: 5 m
- S-Fills mid: 8 m
- S-Fills Far: 16 m

「Close」値は、ステージ上のオブジェクトの深さに対する依存度が少なく、全体的に最良の結果をもたらします。興味深いことに、この値は、スペイシャルフィルシステムとシーンシステムが共有するカバレッジエリア内で、最適なタイムアライメントを実現する値でもあります。

