

L-ISA ハイパーリアル サウンド: なぜ振幅ベースのアルゴリズムがフロントスピーカー構成に 最適なのですか？

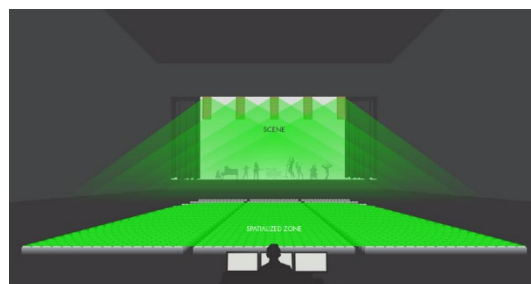


L-ISAテクノロジーは、カバレッジを共有するスピーカーシステムにおいて、**振幅ベースのアルゴリズム**を用いてオーディオオブジェクトをポジショニングします。
これにより、遅延ベースのアルゴリズムと比較して、視覚と聴覚の一貫性を維持しながら、オーディオオブジェクトの自然なセパレーションを最大化します。

L-ISAハイパーリアル サウンド: メリットと条件

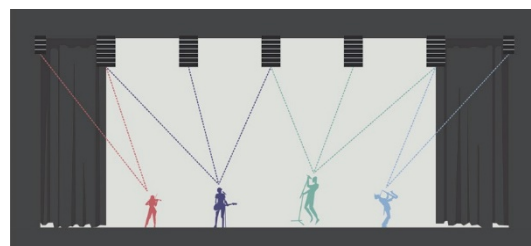
L-ISAハイパーリアルサウンドはライブ環境に高い空間解像度を提供します:

- オーディオオブジェクト*の分離（空間的な非マスクング）を高め、イコライゼーション*とコンプレッション*の使用を限定的なものとしませす。
- パフォーマーの音が「ステージ上のパフォーマーのロケーション」から届くように知覚されることを可能にします。（視聴覚の一貫性）
- ミキシングエンジニアに新たな創造の可能性を提供します。



L-ISAの設計ガイドラインでは、ステージ上方に同じカバレッジ*を共有する最低5つのスピーカーの配置を課しています。

振幅ベースのアプローチによるL-ISAパニングアルゴリズム*は、オーディオオブジェクトの信号を1つまたは複数のスピーカーに送り込むことで、目的とするロケーションへのオーディオオブジェクトの配置を可能にします。このアルゴリズムは、観客全体に視聴覚の一貫性を保った正確なポジション情報を提供するとともに、空間的なマスクングの改善とカラーレーション*の観点から、遅延ベースのアルゴリズムよりもはるかに優れています。

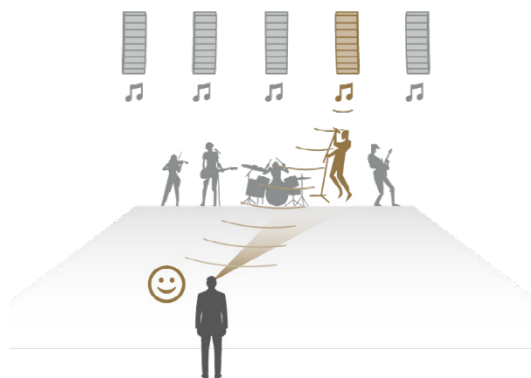


SRにおける空間的な概念

視覚と聴覚の一貫性

観客の正面に設けられたステージで考えると、パフォーマーはステージセットや会場に応じて、ステージの横幅と奥行き、ステージ上または、その上方に沿ってポジショニングされます。このため、視聴覚の一貫性は三次元すべてにおいて想定する必要があります。

垂直方向における視聴覚の一貫性にはスピーカーの設置高が強く関係します。パフォーマーとスピーカーシステム間の垂直角度が30°未満になるようにスピーカーの設置高を設定し、垂直方向における視聴覚の一貫性を確保する必要があります。



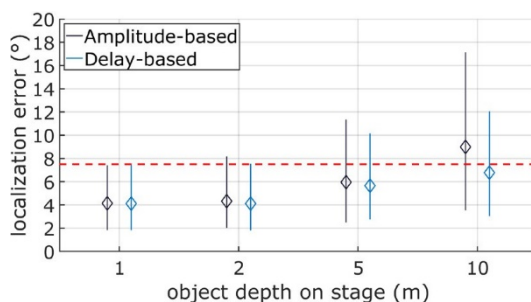
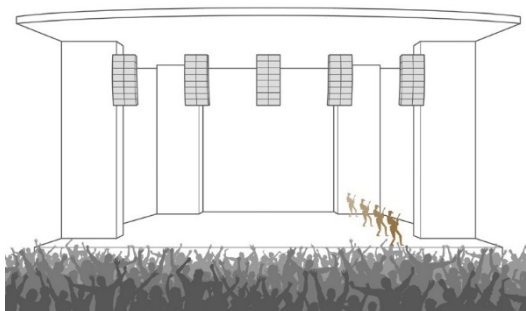
水平方向においては、ターゲットと実際の聴覚位置の間の7.5°までの誤差は、視聴覚の一貫性を保証するためには無視できると見なされます。オフサイドの観客の場合、角度のローカリゼーションはパフォーマンスのステージ横幅方向の位置だけではなく、ステージ奥行き方向の位置にも依存します。

したがって、水平方向における視聴覚の一貫性は以下に依存します：

- ステージ（または会場）の横幅と奥行きに関するパフォーマンス（またはリスナー）のポジション
- スピーカーシステムの分解能（言い換えると：スピーカーの数）
- パンニングのアルゴリズム（振幅ベースまたは遅延ベース）

ここでは、横幅16 mのステージに沿って定期的に配置した5つのスピーカーシステムを検討します。リスナーの耳のレスポンスを演算し、聴覚モデリングによりローカリゼーションを推定する仮想リスニングテストを実施しました。（詳細は付録2を参照してください）

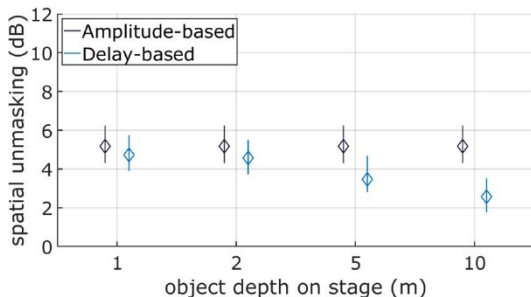
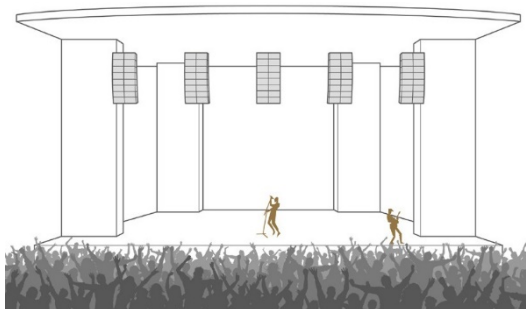
その結果、ほとんどのパフォーマンスとリスナーのロケーションにおいて、水平方向の誤差は7.5°未満であることが示されました。振幅ベース（L-ISA）と遅延ベースのアルゴリズムのどちらも、オブジェクトがステージ奥に位置してもローカリゼーション誤差に関して同様に機能します。（詳細は付録3を参照してください）



複数のオーディオオブジェクトにおける聴覚セパレーションの改善

空間的な非マスキングは、同時に演奏する奏者のなかから「特定の奏者が何を演奏しているのか」を識別するものです。（例えばバックグラウンドボーカルやキーボードに対するリードシンガー）これを高めることにより、リードボーカルの明瞭度を維持するためにバックグラウンドオブジェクトに施すEQやコンプレッションの必要性が低減されます。空間的な非マスキングについて、ここではdB単位で推定します。（詳細は付録2を参照してください）この例では、リードボーカルはステージ中央、バックグラウンドのオブジェクトはステージ中央から上手側6 mに位置します。どちらもステージ上の同じ奥行き方向（スピーカーシステムの後方、1、2、5、10 m）の位置でテストしています。

空間的な非マスキングは、遅延ベースのアルゴリズムよりも振幅ベースのアルゴリズムのほうが常に高い値を示します。遅延ベースのアルゴリズムでは、パフォーマンスがステージ奥へ移動すると、パフォーマンスの見かけ上の角度セパレーションが減少するとともに、ローカリゼーションのぼやけが増加し、結果として空間的な非マスキングの値が半減します。（詳細は付録3を参照してください。）振幅ベースのアルゴリズムは、パフォーマンスの空間的なセパレーションを高く維持し、ローカリゼーションのぼやけを最小限に抑えます。



用語集

オーディオオブジェクト: 空間的なポジション情報などのプロパティが記述されたメタデータを備えるオーディオ入力。

カラーレーション: 自然に発生する音のトーンキャラクターの検知できる変化。

コンプレッション: 信号のダイナミックレンジを低減すること。

カバレッジ: スピーカーシステムが許容可能な周波数応答変動内の直接音を提供するエリア。

イコライゼーション: オーディオシステムの周波数応答を電氣的に調整することを目的としたツールまたはプロセス。

ローカリゼーション: 音の空間的な発生源の知覚。

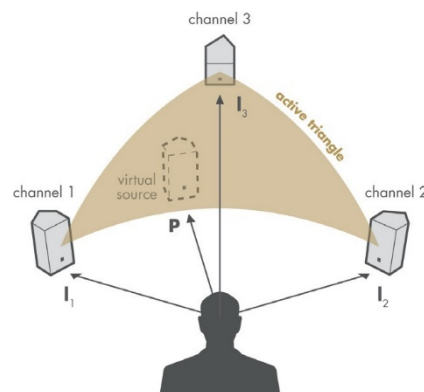
パンニングアルゴリズム: スピーカーに送るオーディオ信号を変更して、関連するオーディオオブジェクトの空間的な発生源を制御するプロセス。

付録 1: 空間パンニング アルゴリズム

振幅ベース: L-ISA パンニング アルゴリズム

L-ISAの空間パンニングは、2D（または3D）パンニングに2つ（または3つ）のスピーカーを使用する振幅パンニングに基づいています。このアルゴリズムは、ベクトルベース振幅パンニング（VBAP）と、多方向振幅パンニング（MDAP）に基づいています。L-ISAアルゴリズムは、VBAP / MDAPアプローチに基づき2つの主要な拡張機能をもたらします：

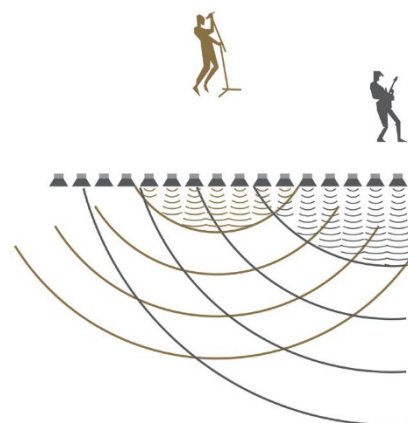
- スピーカー間に配置したオブジェクトの低周波数帯域ビルドアップ補正アルゴリズム
- 複数のスピーカーの組み合わせによるトーンカラーレーションと時間的な影響を最小限に抑え、オブジェクトの横幅を制御するための独自の無相関アルゴリズム



遅延ベース: Wave Field Synthesis

Wave Field Synthesisは、観客全体に向けて一つのオーディオオブジェクトから放射される物理的な音場を理論的に再現する遅延ベースのパンニング技術ですが、これには個別に制御・増幅可能なスピーカーが無数必要となります。

スピーカーの数を減らすと、オーディオオブジェクトのポジショニングが制限されるほか、ほとんどの場合が2次元であり、再現の精度が低周波数帯域のみに限定されます。スピーカーの間隔が大きい典型的な大規模システム（スピーカーの間隔が4 m以上）では、物理的に正確な再現が100 Hz以下の帯域に制限されます。このケースの場合、物理的な波面の再構築はなく、ローカリゼーションはリスナーに最初に到達するスピーカーの寄与によって成り立ちます。（先行ベースのローカリゼーション）したがって、このような少数のスピーカーで構成されたWFSは、デジタルステレオフォニーに類似しています。



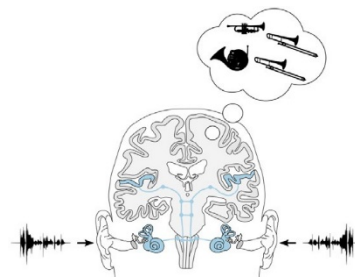
L-ISA と遅延ベースのアルゴリズム

ローカリゼーションに加えて、次の観点からL-ISAと遅延ベースのアルゴリズムを比較することもできます。：

- カバレッジ：遅延ベースのアルゴリズムはアクティブなソースの数を増やすことでカバレッジを拡張しますが、L-ISAスピーカーシステムの設計では、このようなカバレッジの拡張は必要ありません。
- オブジェクトのポジションに応じたカラーレーション：遅延ベースのアルゴリズムはオブジェクトのポジションが変更したときにスピーカー間のタイムアライメントを変更します。これは振幅ベースのアルゴリズムには当てはまりません。
- オブジェクトの移動や軌道によるカラーレーションと影響：遅延ベースのアルゴリズムにおけるリアルタイムでの移動遅延は、高速で移動させるオブジェクトでのクリック音、クロスフェードに十分な時間を与えるための遅延、オブジェクトのポジションによる大幅なカラーレーションの変化など、これらの影響が簡単に発生する可能性があります。

付録 2: パンニング アルゴリズムを評価するフレームワーク

ここでは知覚実験を行う代わりに、スピーカーを用いた空間化システムのローカリゼーション推定の精度が証明されている聴覚モデルを用います。聴覚モデルにはコンサートの状況で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供給し、環境の残響を受けず直接音のみに集中して自由音場の状況をシミュレートするために、無響環境で測定した頭部伝達関数¹を用います。



1 ベルリン工科大学のKemar測定、3mの距離

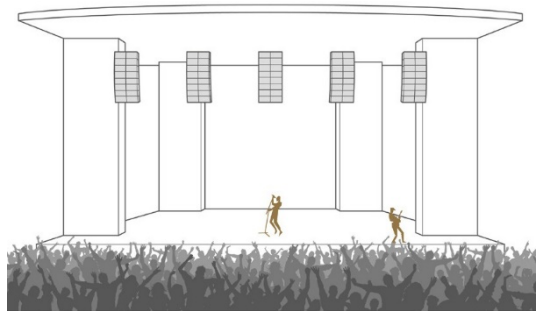
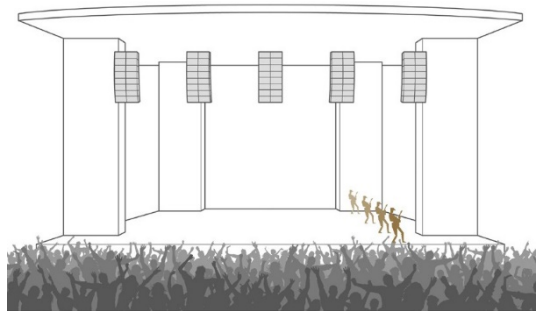
パラメーターはつぎのとおりです:

- 横幅16mのステージに沿って一定間隔(4m)でステージの6m上に配置された5つのスピーカー。
- 81か所のオーディエンスポジション(16 x 16m、2mごと)
- 上手側6mに位置する4つのオブジェクト: スピーカーシステムの1、2、5、10 m後方。

ステージの横幅(12および20m)、オブジェクトの横方向の位置(中心、および中心から外れた)、オーディエンスのサイズ、を変えてテストを実施しましたが、それらの結果は非常に類似し、同じ傾向を示すためここでは紹介しません。

[The Auditory Modeling Toolbox](#) の2つのモデルを使用しています:

- ローカリゼーション推定用のwierstorf2013_estimateazimuthは、水平ローカリゼーションと関連する標準偏差の推定値が得られます。(詳細は付録3を参照してください)
- jelfs2011は空間的な非マスキング用として、ターゲットと干渉源が空間的に分離されている場合のターゲットの音声明瞭度の増加を推定します。テスト状況は、中央に位置するリードシンガーと、上手側に位置するスピーチと同様の周波数レンジを備えた楽器に対応します。



付録 3: ローカリゼーションの精度: 正確さとぼやけ

ローカリゼーションの正確さを評価する場合、参加者はローカリゼーションタスクを実行するように求められ「特定の音をどこで知覚したのか」を回答します。ローカリゼーションテストの結果は、通常2つの要素で分析されますが、そのどちらも、聴覚モデルで対応できます:

- 正確さ: 参加者間の平均回答。
- ぼやけ: 参加者の回答のばらつきに対応するローカリゼーションの不確実性。

ローカリゼーションエラーを評価する場合、すべての回答の標準偏差が推定したローカリゼーションのぼやけに対応するように、参加者の回答のばらつきをシミュレートします。ローカリゼーションエラーは、各テストのリスニング位置について、ターゲットローカリゼーションとシミュレートした参加者の回答の絶対差として算出されます。グラフの値は、ローカリゼーションエラーの中央値(ひし形)と25パーセンタイルと75パーセンタイル(垂直バーの下端と上端)を表しています。

振幅ベースのアルゴリズムは、ほぼ等しいレベルでドライブする**上手外側**の2つのスピーカーのみを用いて上手側のオーディオオブジェクトを作り、約5°のローカリゼーションのぼやけが発生します。遅延ベースのアルゴリズムは、オブジェクトがステージ奥方向へ移動するとローカリゼーションのぼやけの増加が誘発され、通常はこの値が最大で10°まで上昇する可能性があります。(ステージ上10 mのグラフのひし形 中央値) 実際に遅延ベースのアルゴリズムでは、ソースがステージ奥方向に移動するほど、より多くのスピーカーがかなりのレベルでアクティブになり、ローカリゼーションのぼやけが増加します。

ローカリゼーションのぼやけは、ソースをスピーカーの軸上、たとえばステージの**中央**に配置することで最小限に抑えられます。この場合、L-ISAではぼやけが2°まで減少しますが、2°は聴覚システムにおける最低しきい値です。遅延ベースのアルゴリズムでは、ぼやけはオブジェクトの奥行方向の位置とともに徐々に増加し、10 mの距離で8°に達します。この8°という値は、オブジェクトを上手側に位置させたケースと同じレンジと言えます。

