

# リスナーを囲む音楽表現のための楽器音の定位性能の調査

宮家健志

法政大学情報科学部デジタルメディア学科

E-mail: tateshi.miyake.5t@stu.hosei.ac.jp

## Abstract

Although more and more people produce immersive music, its standard is not established yet. In this study, a method of selecting sounds for immersive music is explained. Binaural audio is utilized to listen to immersive audio in headphones. Several tests are held to find the sounds that can be placed in a two-dimensional space as intended. First, a set of HRTFs, head-related transfer function, is chosen for the subject. They are convoluted into samples of various instruments to place them in a two-dimensional space. A subject listens to the audio file and answers the location of the sounds. As a result, the confusion rate for distinguishing between front and back for snares reached 0.125. As for flutes, their front-back confusion was 0.167. On the other hand, violins exhibited a front-back confusion rate of 0.375.

## 1 まえがき

バイノーラル・オーディオを用いたイマーシヴ・ミュージックにおける音の配置について調査する。まず、水平面上での定位を区別できる音を探す。それらを組み合わせて再生した際の定位も検証する。この研究を通して、イマーシヴ・ミュージックを制作する際の定位のガイドを提案する。

近年、メディアでイマーシヴ・オーディオが用いられることがある。イマーシヴ・オーディオは立体音響とも呼ばれ、リスナーを包み込む表現を可能にする。ビデオゲームや映画で使われることが多いが、イマーシヴ・オーディオでミックスされたポップソングも増えてきた。イマーシヴ・オーディオを用いてミックスされた音楽をイマーシヴ・ミュージックと呼ぶ。近年、音楽制作関連のソフトウェアやスピーカーを開発する会社は Dolby Atmos や 360 Reality Audio といった技術に力を入れ始めている。また、消費者は Apple Music や Amazon Music Unlimitedなどを介してイマーシヴ・ミュージックを楽しむ。このように、イマーシヴ・ミュージックの注目が高まっている。しかし、イマーシヴ・オーディオをバイノーラル・オーディオに変換すると音像が頭内に形成されてしまうことがある。現在、消費者はイマーシヴ・ミュージックをヘッドホンやイヤホンで聴取することが多い。そこで、イマーシヴ・オーディオをヘッドホンで再生するためにバイノーラル・オーディオに変換される。バイノーラル・オーディオは音像を頭外に形成できることが特徴である。しかし、音像が頭内に形成されてしまうことがある。すると、イマーシヴ・ミュージック特有の包み込む効果

が薄れる。これを緩和するために背後も含めた定位について調査する。

ポップソングのイマーシヴ・ミュージックの制作を想定する。本研究では音と定位の関係を調べる。現在主流の再生フォーマットであるステレオの楽曲においても、音によって定位の感じ方が異なる。例えば、ベースやキックのような低い音は方向を感じにくい。一方、シンセサイザーのプラックは方向を感じやすい。本研究では、音高と音色の2点に注目する。その結果をもとに、イマーシヴ・ミュージックにおける音の配置のガイドを提案する。例えば前後定位の誤り率が低い音を背後に配置し、前後誤り率が高いものは背後に配置することを避ける。このようなガイドはミキシング・エンジニアがイマーシヴ・オーディオでミックスする際に利益を与える。また、プロデューサーもイマーシヴ・ミュージックに効果的な制作を行えるようになる。例えばステレオで行うミキシングでは、キックやサブベースといった低音はステレオフィールドの中央に置くという慣習がある。本研究は、イマーシヴ・オーディオでのミキシングのガイドとなるスタンダードを確立する助けとなる。

## 2 技術

### 2.1 ポップソングにおけるイマーシヴ・ミュージックの制作

Apple Music や Amazon Music Unlimited など、ストリーミング・サービスでイマーシヴ・ミュージックを聴くことができる。それにより、イマーシヴ・ミュージックへの関心も高まっている。レーベルが楽曲を配信する際は、ステレオだけでなく、イマーシヴ・ミュージックに対応したミックスも配信するようになってきた。

一般的に、録音やアレンジメントはステレオで行われる。オーディオはモノ、もしくはステレオで録音される。サンプルやサンプル・ライブラリもほとんどがモノ、もしくはステレオで収録されている。イマーシヴ・オーディオの制作には Dolby Atmos や 360 Reality Audio が用いられる。例えば Logic Pro や Cubase といった DAW は Dolby Atmos Renderer が標準搭載されている。これを使用することで3次元上に音源を配置できる。ミキシングのためのリスニング環境を整えるためには、スタジオに複数のスピーカーを立体的に配置する必要がある。例えば Dolby Atmos は最低で 5.1.2 (水平面上で囲むように5つ、サブウーファー1つ、上方に2つ) のスピーカーレイアウトを要する。こうして作られたイマーシヴ・ミュージックはヘッドホンやイマーシヴ・オーディオ対応のスピーカーでも再生できる。導入コストと利便性を考えると、スピーカーより一般のヘッドホンの方が手軽にイマーシヴ・ミュージックを楽しむ。ヘッドホンでイマーシヴ・オーディオを再生するた

めにバイノーラル・オーディオが用いられる。

The Kid Laroi 氏と Justin Bieber 氏による楽曲 Stay はイマーシヴ・オーディオでミックスされた楽曲の一例である。ここでは 360 Reality Audio でミックスされたバージョンについて言及する。ステレオのミックスと比較すると、積極的に空間を使用していることがわかる。ボーカルが様々な方向から聞こえてくる。ステレオのミックスでは真ん中と左右から聞こえてくると感じる。それに対し、イマーシヴ・オーディオのミックスでは、正面、左右、左前方、左後方、右前方、右後方から聞こえてくると感じる。特に 1:31 から、2 番のプリコーラスでは左からのボーカルと右からのボーカルがコール・アンド・レスポンスをする。その後メインボーカルが正面から聞こえる。そしてコーラスになるとメインボーカルが右後方や右に配置され、ハーモニーが左に配置されていると感じる。その後も曲の最後にかけて様々な方向からボーカルが聞こえてくる。積極的に空間を使用するイマーシヴ・ミュージックにおいて、楽器を効果的に配置するためのガイドを見つけたい。

## 2.2 イマーシヴ・オーディオ

イマーシヴ・オーディオはリスナーを囲むように音を配置できることが特徴である。定位の扱いが従来のステレオやサラウンドと異なる。サラウンドは channel-based audio と呼ばれる。スピーカー単位で音を操作する。一方、イマーシヴ・オーディオは object-based audio と呼ばれる。背後も含め、リスナーを中心とする立体空間に音源を配置する。それを再生するスピーカーの配置に合わせて変換する。ヘッドホンで再生する際には、バイノーラル・オーディオに変換して再現する。

## 2.3 空間認知

人間が音の方向を認知する仕組みを説明する [1]。水平面上の聞き分けは主に両耳間レベル差 (interaural level difference, ILD) と 両耳間時間差 (interaural time difference, ITD) で行う。例えば音が右前方から来るとき、右耳の方が音量が大きく、左耳の方が音量が小さい。そして、右耳に対して左耳へ音の到達が遅れる。実際には、耳介や肩で回折するため、周波数ごとに異なる ILD や ITD を表現する必要がある。音源から鼓膜、もしくは耳道の入り口までの周波数変化を表した伝達関数を head-related transfer function, HRTF(頭部伝達関数) という。

## 2.4 バイノーラル・オーディオ

イマーシヴ・オーディオをヘッドホンでも実現したい。そこで、バイノーラル・オーディオを用いる。HRTF を反映した音をバイノーラル・オーディオと呼ぶ。本研究では、観察のためにモノ、もしくはステレオの音声をバイノーラル・オーディオに変換する。従来のモノやステレオの音と異なり、HRTF を反映した音像となる。これにより、人間が周囲の音を聴く際と同じ感覚で頭外に音像を形成できる。HRTF はインパルス・レスポンスで計測される。それをモノの音声に畳み込むことで HRTF を反映した音を作ることができる。

バイノーラル・オーディオは、スピーカーでの再生と比べて課題がある。それは、音像が頭内にできることが発生することである。HRTF を用いることで頭外に音像を作ることができる。しかし、それには個人差がある。また、角度や音によって感じる距離が異なる。例えば、ある角度の音像が頭外から聞こえて

も、別の角度の音像は距離が変わってしまうことがある。すなわち、リスナーを中心に均等に音を配置しても、すべての音が意図した角度や距離から聞こえるようにすることは難しい。

## 2.5 HRTF のパーソナライゼーション

HRTF は個人の音の聞こえ方を計測するための方法の一つである。人間の耳や肩の形は個人によって異なる。これにより、人によって音の聞こえ方が異なる。従って、HRTF も人によって異なる。音像定位の性能を上げるためには、HRTF をパーソナライズする必要がある。そこで、本研究では The RIEC Dataset [2] の HRTF を用いる。

The RIEC Dataset は東北大学電気通信研究所の無響室で収録された。105 の被験者やダミー・ヘッドの HRTF を計測したデータセットである。データは SOFA (spatially oriented format for acoustics) というフォーマットで収録されている。被験者に適した HRTF を選出するために、事前にテストを行う。32 種の HRTF をそれぞれ適用したノイズを用いてスイス式トーナメントを実施する [3]。1 秒間ずつのピンクノイズが被験者を公転するようにする。被験者は 2 種の HRTF がそれぞれ畳み込まれたオーディオを聴取し、音像の定位感が良い方を選ぶ。勝ち数の同じもの同士で比較する。そして 2 敗したものは除外する。この工程により、勝ち数が最も多かった HRTF を用いて実験を行う。

HRTF によって定位が異なる。すなわち、音の周波数によって定位の感じ方が異なる。したがって、音楽制作において楽器ごとに定位の感じ方が異なるはずである。その定位感を明らかにしないと効果的なイマーシヴ・ミュージックを制作することは難しい。そこで、楽器ごとの定位感の差を調べる実験を行う。

## 3 実験

### 3.1 音像定位

音高、音色と定位の関係を調べるのが目的である。リスニングテストを行い、意図した定位との差を観察する。

観察に用いる音について説明する。音高はピアノ、ギター、ストリングスピッチカート単音を用いる。オクターブの異なる単音をサンプルライブラリで録音する。2 回ずつ再生する。ピアノは Vienna Symphonic Library Synchron Piano BÖSENDORFER 280VC, ギターは Oragne Tree Samples Evolution Songwriter, ストリングスは 8dio Studio Series Deep Solo Violin と Deep Solo Cello を使用する。様々な楽器を比べて調査する。ピアノ、ギター、シンセリード、ヴァイオリン、トランペット、フルート、キック、スネア、トム、ハイハット、シンバルを用いる。それぞれ 3 種類のフレーズを準備する。ただし、キック、スネア、トム、シンバルはワンショットを使用する。キック、スネア、トムは 8 回、シンバルは 2 回を 4 秒間で均等に並べる。Digital audio workstation (Bitwig Studio 5.1) を使用して並べ、出力する。サンプルは Splice で購入できるサンプルの中から選ぶ。Splice は世界中の音楽プロデューサーが使用するサンプルのストアである。その中から人気が高い順に検索し、なるべくドライなサンプルを選ぶ。3 種類はなるべく奏法、スピード、音色の異なるものを選択する。

次に、これらの音に HRTF を畳み込み、バイノーラル・オーディオに変換する。以下、角度について述べる際は、リ

スナーを中心とした水平面上の回転を指す。正面を 0 度とし、反時計回りを正とする。使用するデータセットは The RIEC HRTF Dataset である。Matlab で sofa ファイルを扱うために SOFAtoolbox を用いる。用意したサンプルに HRTF を積み込むことでバイノーラル・オーディオに変換できる。0 度から 45 度ずつ、315 度までの 8 方向の HRTF を使用する。これにより、8 方向にバイノーラル・パニングしたオーディオを生成できる。

完成したオーディオファイルをランダムに並び替える。Python を用いて一つのオーディオファイルにまとめる。各サンプルの各角度が 1 回ずつ再生されるよう、楽器や角度を混ぜて並べる。各サンプルの再生時間は 4 秒間とする。インターバルは 2 秒間とする。また、順序はテキストファイルとして出力する。

続いてリスニングテストを行い、定位と音の特徴の関係を観察する。ヘッドホンと PC を用いて観察する。使用するヘッドホンは Audio Technica ATH-AVA300 である。これは開放型ヘッドホンである。

まず、被験者にヘッドホンを装着してもらう。初めに、バイノーラル・オーディオの音像が頭外に形成されることを確認する。32 の HRTF を積み込んだ音声から頭の外に音像を形成できる HRTF を選択する。被験者には音像が最も遠く感じるものを、トーナメント形式 [3] で選んでもらう。音像の距離を比べることができない場合、実験を中止する。

続いて楽器の音をランダムに並び替えたオーディオを再生する。まず、聞こえてくる音の方向を 45 度単位で 8 つの角度から選んでもらうことを伝える。オーディオファイルを再生する。被験者はどの方向から音を感じるか選ぶ。結果はサンプルごとに集計する。被験者が選んだ方向をもとに評価する。

### 3.2 コンテキスト上での評価

イマーシブ・ミュージックにおいて意図した位置に定位を感じることができる音の特徴を探ることが目的である。そこで、コンテキスト上での効能を確認するためのリスニングテストを行う。音像定位の実験の結果を考慮した楽曲を制作し、前方だけに配置した曲と比べる。楽曲はポップソングとする。実験の結果より、前後の誤り率が最も低い 2 つの楽器を選出する。そして前後の誤り率が最も低い高さで演奏する。それらの音源を背後に、他の楽器は前方に配置する。背後にある楽器を前方に配置した音源も準備する。これらを被験者に聴いてもらう。音場の広さ、包まれる度合い、楽器の明瞭さ、好み、距離を 9 段階で評価してもらう。

### 3.3 前後誤り率

イマーシブ・ミュージックの強みは立体的に音が配置できることである。ステレオとは異なり、背後にも音を配置することができる。背後の性能を数値化するために前後誤り率 (front-back confusion error rate) を計算する。前後誤り率は以下のように計算される。

$$FBC = \frac{\text{前後誤りの回数}}{\text{試行回数}}$$

被験者が冠状面を挟んで誤った回答をした場合を前後誤りとして数える。ある音の前後誤り率が低いほど前後の分離性能が高いといえる。前後誤り率が 0.750 のとき、最も性能が低い。

0 のとき、前後は完全に分離できている。

### 3.4 メルスペクトログラム

オーディオを分析するにあたり、メルスペクトログラム (mel spectrogram) を使用する。メルスペクトログラムとは、音声の周波数成分をメル周波数スケールに変換したスペクトログラムである。聴覚に基づいた尺度になるため、人間が認識する音の特長を観察できる。

### 3.5 結果

実験結果を表 4 に示す。各楽器に対して行ったテストの回答を集計した。表の左列が楽器、右列がその楽器の前後誤り率である。スネアの前後誤り率が最も低い 0.125、フルートがそれに続き 0.167 であった。一方、ヴァイオリンの前後誤り率が最も高い 0.375、続いて高かったのはリードシンセとキックで 0.333 であった。

表 1 前後誤り率

楽器	前後誤り率
ピアノ	0.292
ギター	0.292
リードシンセ	0.333
ヴァイオリン	0.375
トランペット	0.250
フルート	0.167
キック	0.333
スネア	0.125
トム	0.250
ハイハット	0.292
シンバル	0.250

スネアは打楽器で、アタックが早い。フルートのサンプルにはタンギングが含まれている。一方、ヴァイオリンにはそれが含まれていない。そこで、トランジエントのスペクトログラムを観察した。スネア、フルートのサンプルのうち最も前後誤り率が低かったもののスペクトログラムを順に図??、??に示す。また、ヴァイオリンのサンプルのうち最も前後誤り率が高かったもののスペクトログラムを図 1 に示す。上から順にスネアドラム、フルート、ヴァイオリンである。縦軸が周波数、横軸は時間である。また、橙色であるほど音量が大きく、青であるほど音量が小さい。

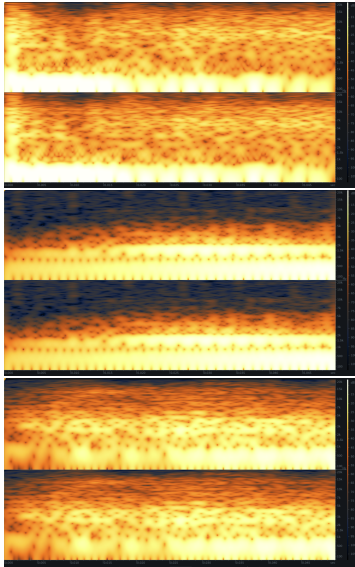


図1 スネア(上) フルート(中) ヴァイオリン(下)のスペクトログラム

続いて、特徴量を得るためにメルスペクトログラムで回帰分析を行った。まず、各サンプルから50ミリ秒のアタックを抽出した。それを31次のフィルタバンクでメルスペクトログラムに変換した。また、それと前後誤り率の関係を調べるために重回帰分析を行った。説明変数は各サンプルのメルスペクトログラムを有する行列、応答変数は各サンプルの前後誤り率のベクトルとした。その結果を図2に示す。係数に対応するフィルタの中心周波数を横軸、その値を縦軸にとった。5次の係数が506 Hzで最も高くなった。

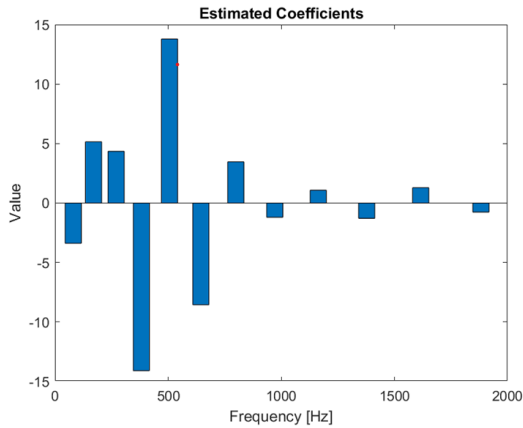


図2 メルスペクトログラムによる前後誤り率の回帰分析

#### 4 考察

イマーシブ・ミュージックを想定し、バイノーラル・オーディオにおける楽器と音像定位について調査した。実験の結果、スネアとフルートの性能が高かった。

前後誤り率の性能を確認するため、ノイズの前後誤り率を調べた。従来の音像定位に関する実験では、ノイズが使用されることが多い。方向を調べるための刺激として、ホワイトノイズ

に HRTF を畳み込み、定位が調べられる。

まず、Matlab を用いて4秒間のホワイトノイズを生成する。それをバイノーラル・オーディオに変換する。これらの音に HRTF を畳み込み、バイノーラル・オーディオに変換する。0度から45度ずつ、315度までの8方向の HRTF を使用する。これにより、8方向にバイノーラル・パンニングしたオーディオを生成する。完成したオーディオファイルをランダムに並び替える。Python を用いて一つのオーディオファイルにまとめる。ノイズの各角度が3回ずつ再生されるようにする。インターバルは2秒間とする。また、順序はテキストファイルとして出力する。続いてリスニングテストを行い、定位と音の特徴の関係を観察する。被験者は3人である。ヘッドホンと PC を用いて観察する。音をランダムに並び替えたオーディオを再生する。被験者はどの方向から音を感じるか選ぶ。結果はサンプルごとに集計する。被験者が選んだ方向をもとに評価する。

実験の結果を表2に示す。表1と比べると、フルートとスネアドラムの性能はこれを上回っている。

表2 ノイズの前後誤り率

楽器	前後誤り率
ノイズ	0.208

フルートとスネアドラムについて調査を進めるため、新たに2人の被験者で実験を行った。各楽器のサンプル数も増やし、前後誤り率が確かに低いことを確認した。まず、ピンクノイズを使用し、頭の外に定位が認められる HRTF を The RIEC HRTF Dataset から選ぶ。観察に用いる音について説明する。フルート、スネアドラム、そしてホワイトノイズを比べて調査する。フルートは9種類のフレーズを準備する。スネアドラムは9種類のワンショットを使用する。8回のヒットを4秒間で均等に並べる。Digital audio workstation (Bitwig Studio 5.1) を使用して並べ、出力する。サンプルは Splice で購入できるサンプルの中から選ぶ。人気が高い順に検索し、なるべくドライなサンプルを選ぶ。9種類はなるべく奏法、スピード、音色の異なるものを選択する。用意したサンプルに HRTF を畳み込み、バイノーラル・オーディオに変換する。0度から45度ずつ、315度までの8方向の HRTF を使用する。これにより、8方向にバイノーラル・パンニングしたオーディオを生成する。完成したオーディオファイルをランダムに並び替える。Python を用いて一つのオーディオファイルにまとめる。各サンプルの各角度が1回ずつ再生されるように並べる。インターバルは2秒間とする。また、順序はテキストファイルとして出力する。続いてリスニングテストを行い、定位と音の特徴の関係を観察する。ヘッドホンと PC を用いて観察する。音をランダムに並び替えたオーディオを再生する。被験者はどの方向から音を感じるか選ぶ。結果はサンプルごとに集計する。被験者が選んだ方向をもとに評価する。

その結果を表3に示す。左列が楽器、右列がその楽器の前後誤り率である。サンプル数を増やし、他の被験者であってもホワイトノイズより高い性能が確認された。

また、回答の傾向についても調べた。それぞれの被験者の回答の角度別前後誤り率を表4に示す。0, 45, 135, 180, 225, 270 はそれぞれ回答した角度を表す。各被験者の135度と225度の

回答を見ると、前後誤り率が 0.152 以下となっている。それに対し、45 度と 315 度の回答では、被験者 B の 315 度を除き、前後誤り率が 0.303 以上となっている。また、各被験者 B と C は、音が後ろから聞こえることが多かったと述べていた。以上のことから、斜めに配置された音は、後ろから聞こえているように感じる人が多いことがわかる。HRTF の影響を観察するため、図 3 に各被験者の HRTF を示す。上から、被験者 A が使用した HRTF、被験者 B が使用した HRTF、被験者 C が使用した HRTF である。横軸は周波数、縦軸は音源に対する相対的な音量変化を表す。また、青い実線は 45 度の HRTF、すなわち左前方の HRTF、橙色の点線は 135 度、すなわち左後方の HRTF である。

表 3 スネアドラムとフルートの前後誤り率

楽器	前後誤り率
フルート	0.250
スネアドラム	0.229
ホワイトノイズ	0.375

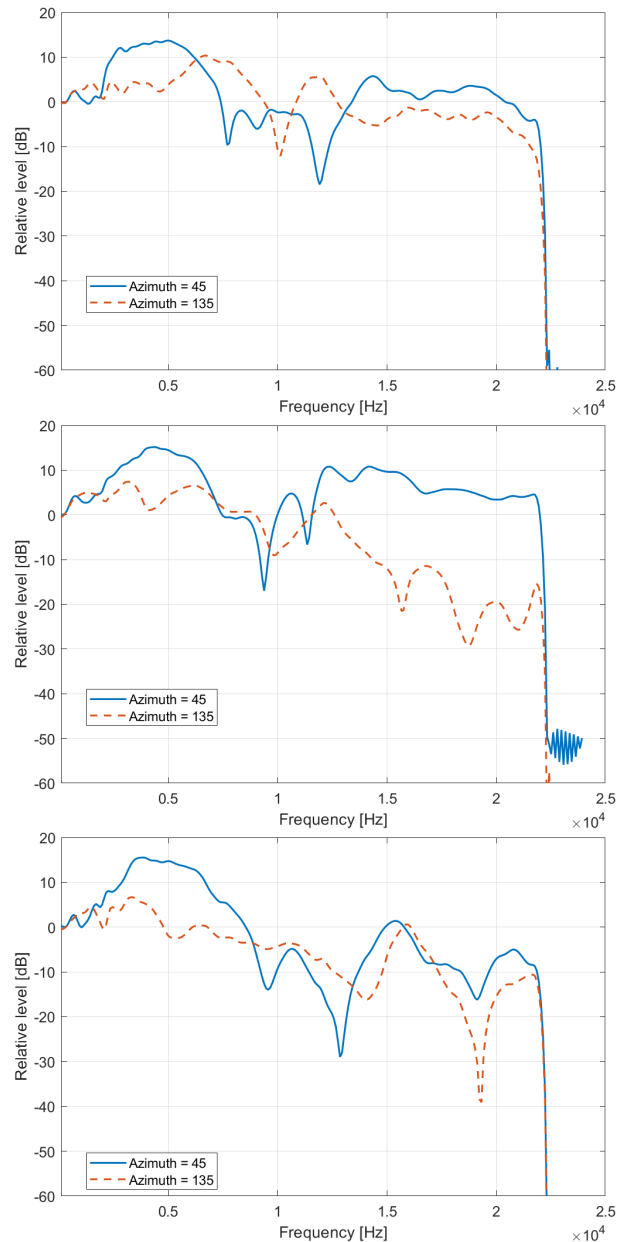


図 3 被験者 A(上) 被験者 B(中) 被験者 C(下) が用いた HRTF

表 4 被験者の回答

被験者	0	45
90	135	180
225	270	315
A	0.273	0.303
-	0.152	0.727
0.030	-	0.727
B	0.842	0.421
-	0.000	0.368
0.000	-	0.000
C	0.947	0.632
-	0.105	0.000
0.053	-	0.737

スネアは打楽器で、アタックが早い。また、フルートのサンプルにはタンギングが含まれている。これらの特徴を踏まえ、トランジェントのスペクトル重心が要因であると推測した。そこで、実験に使用した様々な楽器のサンプルからアタックを抽出した。それらの 5kHz 以上の部分のスペクトル重心を計算した。それらと前後誤り率の散布図を図 4 に示す。横軸がスペクトル重心、縦軸が前後誤り率である。また、赤い点がスネア、黄色い点がフルート、緑の点がヴァイオリンである。相関係数は 0.108 で、相関は見られなかった。

メルスペクトログラムでの回帰分析の結果、506Hz の音が前後誤り率に最も影響していることがわかった。そこで、使用した HRTF を観察した。1000Hz までの HRTF を図 5 に示す。青い実線が 0 度、すなわち正面の HRTF、橙色の点線が 180 度、すなわち真後ろの HRTF である。500Hz 付近を含め、大きな差は見られなかった。



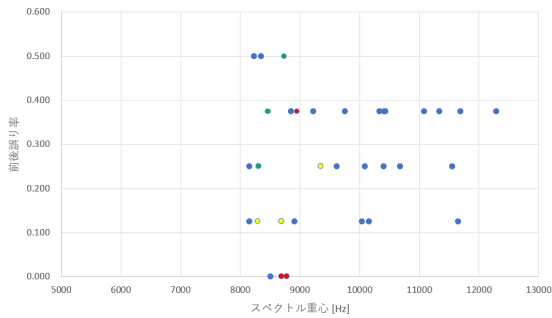


図4 スペクトル重心と前後誤り率の散布図

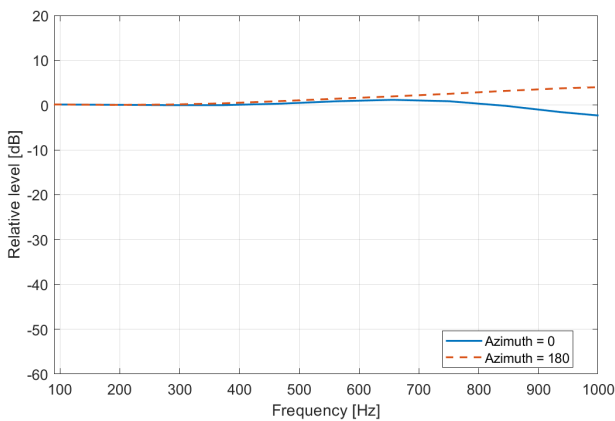


図5 実験に使用した HRTF

A4 を 440Hz とすると、506Hz は B4 と C5 の間になる。そこで、フルートの基本周波数と前後誤り率の関係を調べるために実験を行った。MIDI でフレーズを演奏し、サンプル・ライブラリで録音を行った。まず、digital audio workstation (Cubase Pro 11) にサンプル・ライブラリを読み込む。Straight Ahead Samples Light as a Flutist に含まれるコンサート・フルート (Concert Flute Standard) とアルト・フルート (Alto Flute Standard) を使用する。close mic だけを読み込み、リバーブは使用しないコンサート・フルートを用いて MIDI を演奏する。それを複製して他のオクターブでも録音する。低い音域はアルト・フルートで録音を行う。全部で 3 オクターブ分の演奏を出力する。

続いて、それらをバイノーラル・オーディオに変換し、方向のテストを行う。これらの音に HRTF を畳み込み、バイノーラル・オーディオに変換する。0 度から 45 度ずつ、315 度までの 8 方向の HRTF を使用する。これにより、8 方向にバイノーラル・パンニングしたオーディオを生成する。完成したオーディオファイルをランダムに並び替える。Python を用いて一つのオーディオファイルにまとめる。各サンプルの各角度が 1 回ずつ再生されるように並べる。インターバルは 2 秒間とする。また、順序はテキストファイルとして出力する。続いてリスニングテストを行い、定位と音の特徴の関係を観察する。ヘッドホンと PC を用いて観察する。音をランダムに並び替えたオーディオを再生する。被験者はどの方向から音を感じるか選ぶ。結果はサンプルごとに集計する。被験者が選んだ方向をもとに

評価する。

その結果を図 5 に示す。大きな違いは見られなかった。よって、基本周波数の高さによる影響は観察されなかった。

表 5 フルートの前後誤り率

楽器	前後誤り率
フルート高域	0.250
フルート低域	0.333
アルト・フルート	0.208

## 5 あとがき

本研究では、イマーシブ・ミュージックのための音の配置について調査した。バイノーラル・オーディオを用いてヘッドホンでイマーシブ・ミュージックを聴取する際に、効果的な楽器の配置について考察した。イマーシブ・ミュージックを想定し、バイノーラル・オーディオにおける楽器と音像定位について調査した。実験の結果、スネアとフルートの性能が高かった。しかし、その要因を特定するには更なる調査が必要である。HRTF が人によって異なるように、音の感じ方は人によって異なる。よって、音源定位に関わる音色の特徴を絞り、被験者を増やした実験を行う必要がある。

また、結果をコンテキスト上で評価する。イマーシブ・ミュージックにおいて意図した位置に定位を感じることができる音の特徴を探ることが本研究の目的である。そこで、コンテキスト上での効能を確認するためのリスニングテストを行う。音像定位の実験の結果を考慮した楽曲を制作し、前方だけに配置した曲と比べる。楽曲はポップソングとする。前後の分離性能が高い特徴を持つ音源を背後に、他の楽器は前方に配置する。背後にある楽器を前方に配置した音源も準備する。これらを被験者に聴いてもらう。音場の広さ、包まれる度合い、楽器の明瞭さ、好み、距離を 9 段階で評価してもらう [7]。音の組み合わせによって性能が異なることが予想される。このリスニングテストによってイマーシブ・ミュージックで使用できることを確認する。

## 参考文献

- [1] 岩谷幸雄. 頭部伝達関数による音像定位. 日本音響学会誌. 2017, 73 巻, 3 号, p. 173-180.
- [2] K. Watanabe, Y. Iwaya, Y. Suzuki, S. Takane, and S. Sato, "Dataset of head-related transfer functions measured with a circular loudspeaker array," *Acoust. Sci. & Tech.* 35(3), 159 - 165(2014)
- [3] 岩谷幸雄, 西塔宏二, 鈴木陽一. "勝ち抜き戦方式による個人化頭部伝達関数の選択", 日本音響研究発表会講演論文集. 2004, 3 巻, 2018, p. 559-560.
- [4] J. Kawaura, Y. Suzuki, F. Asano and T. Sone. "Sound localization in headphone reproduction by simulating transfer functions from the sound source to the external ear". *J. Acoust. Soc. Jpn.(E)* 12, 5 (1991)
- [5] A. Kudo, H. Higuchi, H. Hokari, S. Shimada. "Improved method for accurate sound localization". *Acoust. Sci. &*

Tech. 27, 3 (2006)

- [6] H. Wallach. "The Role of Head Movements and Vestibular and Visual Cues in Sound Localization". *Journal of Experimental Psychology*, Oct., 1940, vol.24, no. 4
- [7] Thomas Goerne, Can Karadoğ̃an. "Auditory Scenography in Music Production: Case Study Mixing Classical Turkish Music in Higher Order Ambisonics". 2019 AES International Conference on Immersive and Interactive Audio, March, 2019, Conference Paper 7