

氏名	蘇 恒緯
ヨミガナ	ソ コウイ
学位の種類	博士 (学術)
学位記番号	博音第327号
学位授与年月日	平成31年3月25日
学位論文等題目	(論文) Creation and Perception of Sound Source Width in Binaural Synthesis バイノーラル再生における音像幅の合成及び知覚

論文等審査委員

主査	東京藝術大学	准教授 (音楽学部)	丸井 淳史
副査	東京藝術大学	教授 (音楽学部)	亀川 徹
副査	東京藝術大学	教授 (音楽学部)	西岡 龍彦
副査	東京藝術大学	准教授 (音楽学部)	後藤 英
副査	会津大学	教授	マイケル・コーエン

(論文内容の要旨)

頭部伝達関数(HRTF)の測定により、音源から聴取者の鼓膜まで音の伝達情報が得られ、正確な音源定位を再現できる。そのような、リスナーの両耳に届く音響信号に関係する技術は、「バイノーラル技術」と呼ばれている。バイノーラル技術を用い、立体の空間に音源配置することが可能であり、現実的な聴覚経験を提供することができると期待されている。再現しようとする音場の中に各音源を、HRTF を用いてバイノーラルシンセシスで仮想化することにより、良い空間印象が得られ、柔軟性とインタラクティブの可能性を持つと考えられる。しかし、単純に HRTF と畳み込めば、音の定位情報しか再現できなく、点音源のような狭い音像しか得られない。現実的に存在する音源は単純な点音源ではなく、広さ、大きさを持っている。そのため、より自然な、複雑な空間印象をもたらすため、音像を広げる処理が必要である。

過去の音源の広がり処理に関する研究は、主にスピーカー再生を対象とする。そのため、本研究はバイノーラルシンセシスで音像を広げる方法を考察し、広がり知覚に対する影響を明らかにすることを目的とする。処理方法に関して、周波数帯域を分割し複数の方向に配置するというアプローチを用い、シンセシスしようとする幅の範囲内の方位の HRTF と畳み込めば、各周波数帯域成分を異なる方向に分布できる。そのアプローチをバイノーラルに実装し、音源を広げる処理方法を提案した。以下4つの実験を通し、処理方法の有効性を実証し、信号処理のパラメーターが知覚された音像幅に与える影響、及び音楽制作への応用における全体的なシンセシス効果を考察した。

実験1では、フィルターバンドで音源を1/3オクターブバンドを分け、シンセシスしたい幅の範囲内の各 HRTF と畳み込み、知覚された音像幅にどの影響を与えるかについて聴取実験で調査した。無響室のチェロ、シロホン録音と白色噪音で3種類の音源を用い、5つのシンセシスの幅で刺激を作った。そして、各帯域はどの方向に分布するかを決めるため、4種類の分布方法を使用した。その処理により、空間にどのような音像が知覚されるかを明らかにするため、音像の方位角座標で実験参加者に回答してもらった。結果により、シンセシス幅において音像幅は有意な差が見られなかったが、ヘッドホン再生で頭内定位の問題が生じたため、知覚した音像幅を角度で回答するのが難しいと考えられる。一方、音像の定位を調査するため、回答した音像幅の中心について分析した。其中に2つの分布方法において、音像の定位がずれたと示された。分布方法が音像の定位に強い影響を与えるということを意味した。

実験2では、知覚された音像幅はシンセシス幅によって違うかどうかを考察するため、間接尺度を使用して聴取実験を行った。シェッフェの対比較法を用い、実験1と同じ方法で生成した刺激の音像幅について評価を行った。そして、処理による音質の劣化を調査するため、音色の自然さと空間印象の自然さについても評価された。結果により、チェロ、白色噪音音源は、40度のシンセシス幅以上の場合に、音像幅の広さについて0度の刺激との間に有意差があるが、シロホン音源のほうが音像幅を広げる効果が得られにくいという結果が示された。その原因は、シロホン音源のスペクトル特性が狭いためだと考えられる。また、ランダム分布は、各帯域のパワーによる分布より音像幅が広い傾向がみられる。この結果により、分布方法は音像を広げる効果に影響を与えると示唆された。また、自然さ評価の結果により、音像の広がり処理なしの音源と比べ、処理された音源の音質が有意に劣化していなかったと示した。しかし、分析により評価の個人差があった。実験参加者個人の HRTF と非常に合っていない場合に、音源は明確に定位されず、更に音像の広がり効果を制限すると考えられる。

そのため、実験3では主観聴取の選択による HRTF の個人化によりその問題を改善し、更にシンセシス幅と知覚された幅の関係と他のパラメーターを調査することを目的とする。また、実験1、2の結果により、広がり効果のパフォーマンスは音源のスペクトル特性に依存することと示唆し、周波数帯域の細かさは知覚された幅に影響を与える可能性があると考えられる。処理方法を更に最適化するため、オクターブバンドを更に細かく分割し、そのパラメーターの影響も含めて調査した。そして、先行研究と比較するため、無相関化フィルターによる擬似ステレオの処理方法をバイノーラル再生に実装し、レファレンスとした。聴取実験では音源の広がり音色の自然さについて評価を行った。実験結果により、チェロ音源の場合にシンセシス幅と知覚された幅の間に有意な相関があり、提案された処理方法は有効に音像幅を制御できると示した。また、帯域幅の効果は処理の安定性に影響を与えると示唆された。

実験4では、本研究が提案した処理が、バイノーラル再生向けの音楽制作に応用でき、音像の広がり効果によって空間印象を向上できると検証するため、実証実験を行った。まず、前の実験に用いたアプローチをリアルタイム処理に実装し、プラグインを生成した。普段ミックス作業を行なっている実験参加者2名に、そのプラグインを DAW で使用してサウンドデザインを行ない、映像と合わせて各効果音の定位と幅をリアルタイムで調整した。そのミックスを用い、主観評価実験を行い、実験参加者に効果音の全体的な空間印象について評価してもらった。その結果を分析し、音像の広がり処理により空間印象を向上できると検証した。

以上の調査、検証実験を通して、本研究で提案された処理方法は、適切な音源、パラメーターの条件の下に、有効にモノラル音源の音像幅を合成と制御することができるし、オーディオ制作へも実際に応用できると検証した。しかし、個人差と音源特性への依存性があり、今後の研究でさらに調査と改善する必要があると意味した。

(総合審査結果の要旨)

これまで映像作品や音楽を再生する方式として、2台のスピーカを視聴位置前方に配置したステレオ方式に対し、5台のスピーカで視聴位置を囲むサラウンド方式が提案され、さらに近年では視聴位置を中心にしてスピーカを上方・下方にも配置する三次元音響再生システムが映画館やテレビ放送などで導入されている。より多くのスピーカを用いることで、空間内の音像の定位や大きさの表現についての自由度や精度が向上することが明らかになっており、また、それらの空間表現をヘッドホンを用いて再現しようという試みも開発・実用化されつつある。

こうしたシステムでは、音信号（オブジェクト）と定位など音に付随する情報（メタデータ）を分離し再生時に合成を行うオブジェクト方式と呼ばれる手法が用いられることも多い。メタデータは「音響定義モデル（Audio Definition Model）」などによって、定位、音像の大きさ、残響の量などが指定されることになっているものの、その中でもヘッドホン再生時にモノラル音源から音像の大きさ（音像幅）を作り出す方法については決定打がないのが現状である。それを受け、本研究ではヘッドホン再生において聴感上の音の大きさを調整する新たな手法を提案した。具体的には、音像幅を広げたい元音源を細かい周波数帯域ごとに分割し、それぞれの帯域に異なる方向の頭部伝達関数をかけてから再合成するというものである。ここで、分割帯域数の設定、頭部伝達関数の適用方法、再合成の方法、頭部伝達関数の個人化など、いくつもの変数が複雑に絡み合っている。そこで、申請者は聴取実験を通して提案手法の精度を向上させるとともに、その有効性の客観的な確認を行っていったのが本研究の骨子である。

上記のように本研究は内容も手法も極めて工学的であるが、提案手法を録音編集ソフトウェアの中で利用できるようにし実際に作品制作経験のあるエンジニアに制作現場に近い設定で使用してもらうことで有効性の確認を行ったことにより、本研究を単に理論上のものではなく実用性までも視野に入れたものにできた。結果として、手法の提案と検証にとどまらず、音色印象を犠牲にすることなくリアルタイムで音像幅の拡大が可能となるシステムの実装までを行う充実した研究となった。

論文審査会では、実験に用いた元音源の特性や、実験参加者の耳形状の個人差などについて問題提起がなされたものの、研究の新規性・有用性についての評価は審査員が一致して認めるところであった。本論文は博士の学位にふさわしいものである。