

音源特性を三次元空間内に提示する VR ミキシングシステムの提案

2331067 近藤建樹 成見研究室

1 はじめに

近年、アマチュア音楽家の作品がインターネット上で広く知られることも珍しくなくなり、音楽制作を志す者が増えてきた。しかし、初心者が音楽制作を技術として習得することには大きな壁がある。音楽制作を行うにあたっては、自らの聴覚を頼りに音を分析し、アレンジを加えていくことが必要であるが、聴覚に頼った技術の習得には個人差はあるものの一定の期間を要することが多い。

一方で、聴覚を視覚によって補助する試みは多く行われてきた。現代ではディスプレイを介して多様な音情報が視覚的に表現できる。しかし、VR デバイスを用いて音声情報を表示する試みはほとんど行われていない。ただし、技術の進歩によって今後 VR デバイスが音楽制作にも取り入れることが出来るようになることが予想される。

そこで本研究では、立体感を人間が感じ取りやすいという VR デバイスの特性を活用し、新しい音楽制作ツールの提案を行う。音声情報を表示するための円筒座標空間である Cylindrical Acoustic Space(CAS) と、CAS 内で音声情報を表示する音声視覚化帯の 2 つを用いた新しい音声情報の表示方法を提唱する。特に音楽制作の工程に含まれるミキシングに注目し、視覚情報を活用して直観的にミキシングを行う事が出来るシステムの開発を行う。また、開発したシステムを使用してミキシングを行った際に、使用者が感じる作業負荷や作業結果にどのような差異が生じるかを既存の手法でミキシングを行った際と比較して評価した。この結果、音楽制作に VR デバイスを導入することによってミキシング作業の心理的負荷が小さくなる利点があることが分かった、音楽制作において本システムを用いることで、聴覚に頼らずとも多くの人が楽しむことが出来る可能性がある。

2 ミキシングとは

音楽制作におけるミキシングとは、音を混ぜ合わせることにほかならず、記録された複数の楽器の音をもとにして 1 つの出力音を作成する。ミキシングを行う目的には大きく分けて芸術的側面と技術的側面がある。芸術的側面では音の鳴る空間の演出などが行われるが、本研究では後者の技術的側面を扱う。特に複数の音同士が互いに影響しあうことで、人間が音を聴きとりにくくなってしまいう「マスキング」を回避する作業に注目する。

マスキングは哺乳類の聴覚特性により発生する現象であ

る。人間は同じ方向から聴こえてくる同じ周波数帯の音を別個のものとして知覚することが難しいという性質を持っている。そのため、複数の楽器音を混ぜ合わせる際には、同じ周波数帯で鳴る音を同じ位置に配置しない、または、周波数帯を変更し互いに影響が起きにくいようにすることが求められる。

3 提案システム

立っているユーザを縦軸とし、軸の上方向を周波数、周方向を音源の方向、径方向を音圧とする円筒座標空間である、Cylindrical Acoustic Space(CAS) を提案する。また、CAS において音声情報を 3D 空間に提示する方法として、ある音の周波数 f 毎の音圧 v_f 、音の広がり w_f 、音の方向 p_f を表すと、図 1 のような帯状の開曲面となる。この開曲面を以後音声視覚化帯と呼ぶこととする。

開発したシステムは図 2 のように、PC 端末で起動された音楽制作ソフトウェア (DAW) 上でプラグインとして動作する VST プラグイン部と、VR 端末で動作する VR アプリケーション部から構成される。この 2 つの端末間で相互に通信しあうことで本システムは動作する。

VST プラグイン部では音声の解析を行い解析結果を VR アプリケーション部に送信する。一方、VR アプリケーション部では解析結果を受信し、HMD に解析結果を表示する。また、ユーザが行った操作を VST プラグイン部に送信する。そして、VST プラグイン部は操作結果をもとにして音声にリアルタイムで加工を行う。

4 評価

一定の視力、聴力を有する被験者 24 名に対して実験を行った。実験は、音声視覚化帯の個人ごとの広がり

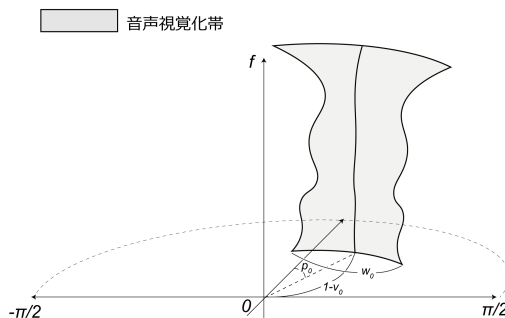


図 1 CAS 内における音声視覚化帯の表示

を行う実験、提案手法と既存手法の比較実験の二種類を行い、前半のみの参加者が1名いるため、その他の実験・アンケートへの参加者は23名となっている。実験は「音源特性を三次元空間内に提示するVRミキシングシステム」(H24073)として電気通信大学の人を対象とする研究に関する倫理委員会で承認後実施した。

提案手法と既存手法の比較実験では、NASA-TLX[1]日本語版を用いた主観的メンタルワークロード測定を行った。その結果図3に示すように提案手法を用いることで、Mental Demand(知的・知覚的欲求)、Temporal Demand(タイムプレッシャー)、Frustration(フラストレーション)の指標において提案手法が下回る結果となった。また、Overallでも提案手法が下回り、両側t検定を実施したところ、 $p = 0.030 < 0.05$ となった。よって、本システムがミキシングタスクのユーザへの心理的な負荷を有意に軽減することが示された。

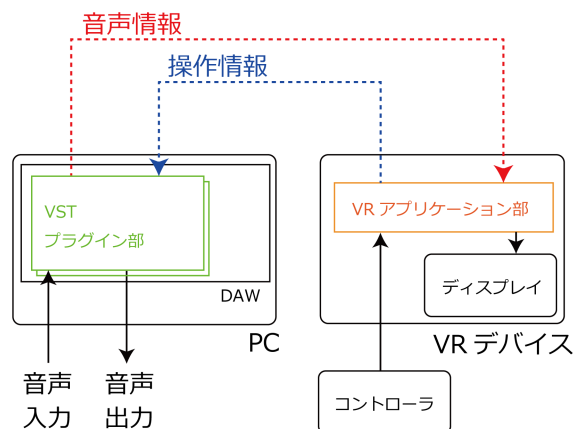


図2 システムの概要

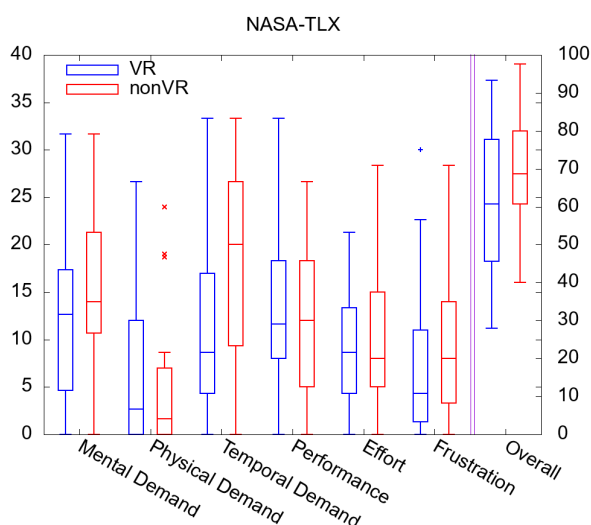


図3 NASA-TLX 日本語版による評価結果

5 関連研究

Anders Bargum ら [2] は、VRを用いたミキシングツールの実装を行い、音楽制作ソフトウェアと統合するためのフレームワークを作成した。また、それを既存手法と比較し、VRの感覚的な利点と分かりやすい定位感が3Dと立体音響のミキシング工程を改善できることを示した。この研究はVRを用いてミキシングを行う試みという点で本研究と類似しているものといえる。しかし、この研究は音楽制作に焦点を絞ったものでは無く、本研究におけるCASのような周波数特性を視覚的に表現する工夫はなされていない。

岡ら [3] はユーザによるインタラクションが可能なVRオーディオビジュアライザを作成し、その妥当性についての評価を行った。この研究は音の周波数特性をVR空間内に表現し、ユーザにその周波数特性を操作させるものであるという点において本研究と類似している。しかし、この研究では複数の音源に対するミキシングについて焦点を当てたものではない。

6 おわりに

本研究では、VRの特性を活用した新しい音楽制作ツールを開発した。また、音声情報を表示するためのCASと音声視覚化帯を提案した。これにより、音楽制作初心者でも直観的にミキシングを行う事が出来るようになった。

さらに、開発したシステムと既存手法を用いて23名の被験者にミキシングタスクの実験を行った。NASA-TLX日本語版を用いて主観的メンタルワークロードに関する評価を行ったところ、知的・知覚的欲求、タイムプレッシャー、フラストレーションの指標において提案手法が既存手法を下回り、総合的にも有意に提案手法の心理的な負荷が小さいことが分かった。

参考文献

- [1] NASA Task Load Index. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. "https://humansystems.arc.nasa.gov/groups/TLX/" (参照 2025/1/20)
- [2] Anders Riddershom Bargum, Oddur Ingi Kristjánsson, Péter Babó, Rasmus Eske Waage Nielsen, Simon Rostami Mosen, and Stefania Serafin. Spatial Audio Mixing in Virtual Reality. *Sonic Interactions in Virtual Environments*. pp269–302. (2022)
- [3] 岡あゆみ, 栗原一貴. VRでの音の可視化とインタラクションの考察. 「エンタテインメントコンピューティングシンポジウム (EC2021)」2021年8月.