

論文要旨

ニューラルネットワークによる音色の自動変換

学際科学科 総合情報学コース

08-192021

陶山大輝

指導教員 金子知適

音楽は世界中で楽しまれている。音楽の作成方法には様々なものがあり、既存の曲をアレンジして新しい曲を作成する Remix と呼ばれる方法がある。しかし、Remix は初心者が手軽に行えるものではない。そこで、コンピュータプログラムによる補助が役に立つと考えられる。本研究では、Remix の代表的な方法の一つである音色の変換に着目し、プログラムによる変換手法を提案する。

音色の変換を行うためには、ある楽器の音を異なる楽器の音へ変換する技術が必要である。そこで、本研究では Pix2pix [1] を音色の変換に応用した。Pix2pix はニューラルネットワークにより自然な画像を生成する手法である Generative Adversarial Networks (GAN) [2] を用いて画像のスタイル変換を行う手法である。

本研究では、ギターの単音をハーブの単音に提案手法を用いて変換する実験を行った。その結果、音の大きさが変わることなどの問題はあったものの、ほとんどの音で音程を維持したまま音色の変換を行うことに成功した。また、データセットの一部の音のみで学習を行った場合でも、ほとんどの音で音程を維持したまま変換を行うことができた。

- [1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2018.
- [2] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.