

2016年度 卒業論文

word2vec を用いた n-gram モデルによる楽曲類似度解析

龍谷大学 理工学部 数理情報学科
T130100 和田 菜子
担当教員 佐野 彰

概要

楽曲は、漫画や小説のように要約することが難しく、自らが試聴して好みの楽曲であるか判断するしかない。したがって、数多くの楽曲の中から求める楽曲を探すためには、多くの時間や労力を費やさなければならない。

そこで、楽曲のメロディラインから word2vec を用いて楽曲の特徴ベクトルを作り、楽曲類似度を求める研究を行う。word2vec とは自然言語処理の手法で、単語の特徴ベクトルを得るための手法である。単語の特徴ベクトルは2層のニューラルネットワークによって学習される。word2vec では、単語とその周辺にある単語を用いて単語の共起確率を計算する。十分なデータが与えられれば、ある単語の意味と類似している単語の推測を高い精度で行うことができる。これにより、ある単語の周辺にはどの単語があるときに登場しやすいかが学習される。

word2vec は言語処理の分野の手法なので、単語を持たない楽曲にそのまま適用することができない。したがって、楽曲に word2vec を適用するために、MIDI 形式の楽曲データを MML 形式に変換し、MML 形式の楽曲データから音符テキストデータを作る処理を行った。さらに、音符テキストデータから隣接する n 個の音符を1つの音声単語とすること (n -gram) で、楽曲を word2vec により処理できるようにした。

まず、多くの楽曲で比較する前に少数の曲で解析を行った。MML 形式の楽曲データを n -gram での分かち書きを行い、音声単語を生成した。次に、word2vec を用いて n -gram に分けた音声単語の特徴ベクトルを得た。得られた特徴ベクトルを川野によって考案された楽曲間の類似尺度を用いて計算を行った。得られた計算結果を MDS によって2次元に図示することで、類似尺度の特徴や、 n -gram で区切った音声単語の特徴を得ることができた。

次に、先の実験で扱った3曲を含む10曲に曲数を増やし、楽曲間の類似度を測り、MDS で図示した。実験には J-pop、演歌、ロックの3つのジャンルの楽曲を使用した。それぞれのジャンル間の特徴や、3曲の類似尺度の変化を結果で獲得した。実験の結果、ジャンルによる楽曲間の分離や、 n -gram で区切った音声単語による楽曲間の尺度の変化を確認することができた。しかし、類似尺度の間の距離が近くても実際に曲を聴くと類似しているとは思えない楽曲もあったため、word2vec の学習方法を変更することや、楽曲数を増やすことが課題である。