

楽曲データからのスケール自動判別

赤澤 豪樹

近年、膨大な楽曲データを取り扱う音楽配信サービスが多くのユーザを獲得している。デジタルコンテンツ市場全体の中でも音楽コンテンツは大きな役割を果たしている。音楽コンテンツに限らず、この種のサービスにおいては一部のものが極端に高い人気を誇り、売上を伸ばす傾向が強い。よって、楽曲データの推薦において人気や売上だけを考慮していたのでは、適切な推薦が困難となる。

こうした状況のなかで、大量の楽曲データを効率良く解析し、ユーザが求める楽曲を幅広く提供できるような楽曲推薦手法の重要性が増している。そのためには、より深く楽曲の構成要素を解析し、ユーザが好む要因を発見することが良い推薦の第一歩であり、その一役を担うものがスケール（音階）解析であると考えられる。スケールとは、ある制約をもって並べられた音の列であり、主に旋律や伴奏を形作る源となる。明示的にせよ暗示的にせよ、大抵の楽曲は何らかのスケールを基にしている。また、スケールは楽曲の雰囲気や性質に大きな影響を与えるものであり、特徴を表す一つの要素となる。しかし、現時点では、コード判別などを実践している研究はあるが、詳細なスケール解析を行っているものは存在しない。また、人手によってスケールを判別する方法はあるが、時間と労力がかかる上に、大量の楽曲に対してそれを行うのは困難である。さらに、その場合主観的な判断に左右されてしまう可能性も高い。

そこで本研究では、機械学習を基に楽曲におけるスケールを自動的に判別する手法について考察する。機械学習手法としては RBF (Radial Basic Function) カーネルを応用した SVM を用いる。スケール構成音との一致率を算出する独自の手法に基づきラベル付与を行い、学習データを作成した後、それらを用いて学習を実行する。生成された推定モデルが、未知の楽曲のスケール名を推定する。

評価実験として、スケール教則本に紹介されているフレーズを用い、推定モデルの精度を検証した。まず、スケールとの一致率の基準を示す実験、極端にデータ数の偏った不均衡データに対する手法の検証を行った。一致率の基準値として、スケール構成音に対して 9 割程度とした場合に最も高い正解率が得られることが判明した。また、不均衡データに対する手法として、オーバーサンプリングとアンダーサンプリングとを組み合わせた手法が有効である結果が示された。さらに、学習データ作成において、音の出現回数のみに着目したデータと音の長さまで取り入れたデータの二つを用意し、それぞれ正解率の検証を行った。その結果、音の長さを取り入れたデータの方が、概ね正解率が向上することがわかった。これは、音に対する重み付けがより明確となり、重要度の低い音と高い音が区別しやすくなったためと考えられる。

(指導教員 鈴木伸崇)