

分散和音の影響に着目した和音変化検出手法の改善

土本 慎史[†] 西垣 貴央[†] 小野田 崇[†]

青山学院大学大学院 理工学研究科[†]

1. はじめに

自動採譜や楽曲推薦などの音楽情報を扱う研究において、wav データのような音楽データから和音の変化したところを検出する和音変化検出技術は、多くの応用に用いられ非常に重要である。従来研究では、同時和音同士の距離を測る調性空間の類似距離の極大を使う方法[1]や、打楽器音を分離し、離散フーリエ変換を用いて拡張した調性空間の類似距離の極大を使う方法[2]が提案されている。また、MIDI データでは、分散和音を考慮した SVM による和音変化検出手法[3]が提案されている。一般的に和音は同時に鳴る数の異なる音高のことを表すが、和音にはその複数の音高を順に鳴らす分散和音というものも存在する。これまでの wav データを扱った研究では、この分散和音について考慮されていないため、和音変化の誤検出が増えてしまっていた。そこで本稿では、wav データに対して分散和音の影響を考慮した和音変化検出手法の提案をする。

2. 提案手法

音楽音響信号から和音の変化位置を検出する問題に取り組む。図 1 に提案手法の概要を示す。

2.1 前処理

入力された音楽音響信号からメディアンフィルタを用いて打楽器音の分離を行い、Robust-PCA で歌声音の分離を行う。その後、打楽器音と歌声音が分離された信号から、非負最小二乗 (nnls) によって 12 音階を表す 12 次元のクロマベクトルを算出する。

2.2 拡張した調性空間に写像[2]

2.1 で求めたクロマベクトルを拡張された調性空間に写像し、調性空間ベクトルを得る。

2.3 分散和音の影響を考慮[3]

分散和音を考慮するには、考慮するベクトルに幅が必要になる。最初に、調性空間ベクトルの自己類似度行列 D を求める。行列 D の対角線を横切って、ある区間 i とその次の区間 $i + 1$ の新規性 nov を以下の式で求める。

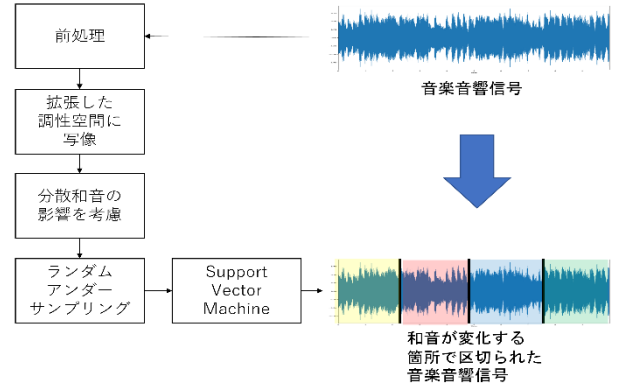


図 1 提案手法の概要

$$nov_{i,i+1} = \sum_{m=1}^L \sum_{n=1}^L C_{m,n} D_{i-L/2+m,i-L/2+n}$$

$$C_L = \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} \otimes O_{L/2}$$

$O_{L/2}$ は $(L/2) \times (L/2)$ の単位行列である。 C_L は、サイズ $L \times L$ 及び標準偏差 $\sigma_g = 0.5$ のガウスカーネルとの積により、ガウスチェッカーボードとなる。この nov が高い場合、和音が変わった可能性が高い。 nov を $L = \{4, 8, 16, 32\}$ でそれぞれ算出する。 L が小さいと細かい変化に敏感で、 L が大きいと正確性が高い。また、各 L のある区間 i の特徴量を $nov_{i-3,i-2} \sim nov_{i+3,i+4}$ まで拡張する。これによって 28 次元の特徴量を得る。

2.4 ランダムアンダーサンプリング

一般的な音楽は和音が変わったクラスのデータと和音が変わっていないクラスのデータの数の比は大きく異なる。[3]の手法ではクラス間のデータ数の比を考慮しておらず、wav データに対して使用した場合、多くのデータを多数クラスである和音が変わっていないクラスと識別してしまう。そのため、学習データ数のアンバランスさを考慮する必要がある。学習データ数のアンバランスさを考慮する手法としてランダムアンダーサンプリングが存在する[4]。ランダムアンダーサンプリングは、ランダムに多数クラスの学習データを削減する手法である。このランダムアンダーサンプリングによって、多数クラスである和音が変わっていないデータを削減する。

2.5 Support Vector Machine

ある区間 i ごとに和音が変わったか変化していないかの 2 値分類を、SVM (Support Vector

Improvement of over-detection of chord change detection method focusing on the influence on broken chord
[†] Shinji Tsuchimoto, Takahiro Nishigaki, and Takashi Onoda, Aoyama Gakuin University Graduate School of Science and Engineering.

Machine) を用いて行う。SVM は、機械学習手法の一つで、2 値分類問題を解くためのアルゴリズムである[5]。SVM は他の機械学習手法に比べ少ない学習データでも安定した汎化能力を持つとされている[5]。[3]のSVMの学習では、和音が変化した個所とその誤差分である前後3区間分を和音が変わったクラスとしている。本手法では和音が変わった個所のみ和音が変わったクラスとして学習させ、その誤差分である前後3区間分は両クラスの学習から省く。またSVMの出力では、複数の区間を連続で和音が変わったと識別する可能性がある。その場合は連続した区間の平均位置を使用する。

3. 評価実験

3.1 実験条件

関連研究[2]と同じThe Beatlesの楽曲110曲のwavデータと、正解データ[6]を用いて5分割交差検証を行った。

アンダーサンプリングでは、和音が変わったときのデータと和音が変わっていないときのデータ数の比を1:1.5になるように、多数クラスのデータを削減した。

SVM は rbf カーネルを用いた。また、 $C: \{1000, 5000, 10000\}$, $\gamma: \{0.005, 0.01, 0.05, 0.1, 0.5, 1\}$ の範囲をグリッドサーチで求めた。グリッドサーチは学習データを4分割して行っている。

評価指標として、再現率、適合率及びF値を使用する。関連研究[1][2]と同じく、検出の評価時に誤差として ± 0.278 秒内の検出は正解とする。

3.2 実験結果

実験結果を表1に示す。関連研究は機械学習手法を用いていないため、学習率は空白となっている。関連研究に比べ提案手法はF値が約3ポイント、適合率が約13ポイント向上した。適合率は、検出結果が本当に正しい割合である。そのため、正確な和音変化を検出できていると考えられる。また、F値は再現率と適合率の調和平均であるため、再現率と適合率のバランスが良くなった提案手法のF値が向上したと考えられる。分散和音が含まれる楽曲の分散和音が特徴的な部分に対する検出結果の例を図2に示す。図2では、青線は和音が変わったとして検出した個所で、橙線は正解和音変化箇所とその誤差を表す縦軸が1の時和音変化力所、0.5の時和音変化力所の誤差範囲、0の時に和音変化力所でないことを表している。橙線内

表1 実験結果

手法	学習率	再現率	適合率	F値
関連[2]		81.90%	62.80%	69.10%
提案	90.60%	72.00%	75.90%	72.20%

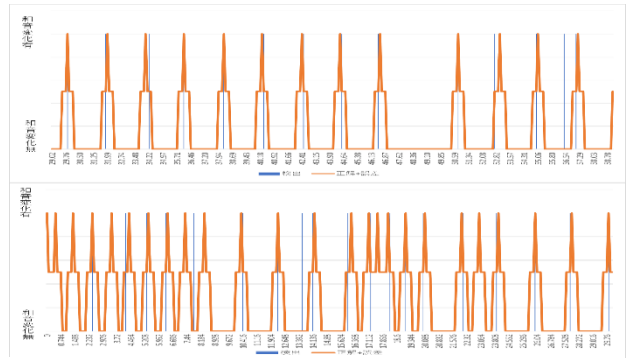


図2 実験結果例 (And I Love Her (上), Till There Was You (下))

に青線が含まれている場合、正しい検出ができているとなる。分散和音が含まれる楽曲に対し、効果的な和音の変化検出ができています。また、関連手法[2]に比べ再現率が約10ポイント減少した。これは、SVMの出力で和音が変わったと識別した区間が連続した場合、平均位置を取るが、短いスパンで複数の和音が変わっている場合、1つの和音が変わった区間として扱われてしまうため、検出数が少なくなっていることが原因と考えられる。この問題の解消は今後の課題である。

4. おわりに

本稿では、分散和音の影響に着目した和音変化の検出手法について述べた。実験結果では、提案手法が分散和音の影響によって発生する誤った検出を減らすことが確認された。今回、データのアンバランスさを考慮する手法としてランダムアンダーサンプリングを使用した。オーバーサンプリングやそれらの併用も有効であると考えられる。また、本手法ではパラメータとして調性空間や新規性のサイズなどのある特定の値を使用した。楽曲によって適した値があると考えられる。従って後は、楽曲それぞれに適した和音の変化検出手法の考案を目指す。

参考文献

- [1] Christopher Harte, Mark Sandler, Martin Gasser; Detecting Harmonic Change In Musical Audio, Proceedings of the 1st ACM workshop on Audio and music computing multimedia, 2006, pp21-26.
- [2] Pedro Ramoneda Franco, Gilberto Bernardes; Harmonic Change Detection from Musical Audio, Faculty of Engineering of the University of Porto Master's thesis, 2020.
- [3] Joshua Morman, Lawrence Rabiner; A System For the Automatic Segmentation and Classification of Chord Sequences, Proceedings of the 1st ACM workshop on Audio and music computing multimedia, 2006, pp.1-10.
- [4] Yuchun Tang, Yan-Qing Zhang, Nitesh V, Chawla and Sven Krasser; SVMs Modeling for Highly Imbalanced Classification, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B: CYBERNETICS, VOL. 39, NO. 1, 2009.
- [5] 小野田崇, サポートベクターマシン, オーム社, 2007.
- [6] Isophonics Datasets; <http://www.isophonics.net/datasets>.