

# 楽器の階層的分類を考慮した SVM による音源同定

## Sound Source Identification Using SVM with Hierarchical Musical Instrument Classification

北原聡志 甲藤二郎  
Satoshi KITAHARA Jiro KATTO

早稲田大学大学院 理工学研究所  
Graduate School of Science and Engineering, Waseda University

### 1. まえがき

音源同定はパターン認識の応用と捉えられる。先行研究例として、GMM を用いて楽器毎のモデルを作成する手法があり、5 種類の楽器に対して平均 70[%]程度の認識率が報告されている[1]。我々は一検討として Support Vector Machine(SVM)を用いて、楽器毎に、認識対象の楽器とその他の楽器の特徴量をそれぞれ学習させた 2 値分類器を作成し、楽器認識を行う手法を提案した[2]。しかし、この手法ではどの SVM も反応しない場合、認識不可能という問題があった。これを解決するために、本稿では楽器の階層的分類に基づく SVM を用いた手法について述べる。

### 2. 提案手法

#### 2.1 楽器同定処理

楽器の階層分類を考慮し、SVM の基本原理である 2 値クラス分類を繰り返していくことで楽器同定を実現する(図 1)。研究対象の 6 種類の楽器は大別すると減衰系(Cg, Eb, Pf)と持続系(Cl, Tr, Vn)から成る。減衰系の中では、Cg と Eb は構造や演奏方法が似ていることから 1 つのカテゴリーと考えると、Cg&Eb のクラスと Pf のクラスに分けることができる。また持続系では Cl と Tr が管楽器、Vn が弦楽器に属する。以上のことを踏まえて識別器を作成する。例えば、図 1 における減衰系 持続系分類は、一方のクラスに減衰系楽器の特徴量を、もう一方のクラスに持続系楽器の特徴量を学習させた SVM を用いることで行う。

#### 2.2 ケプストラムの利用

楽器特徴量として FFT ベースのケプストラムを利用する。FFT は窓幅 4096 点とし、1/2 ずつオーバーラップさせながら解析を行う。また、一次差分 ケプストラムおよび二次差分<sup>2</sup>ケプストラムをそれぞれ計算し、特徴量に当てる。特徴量次元数は各々26、計 78 次元とする。

#### 2.3 MFCC の利用

MFCC を用いて楽器認識を行い、ケプストラムとの結果を比較する。MFCC はフィルタバンク分析により計算する。次元数はケプストラムと同様 26 とし、一次差分 MFCC、二次差分<sup>2</sup>MFCC を加え、計 78 次元とする。

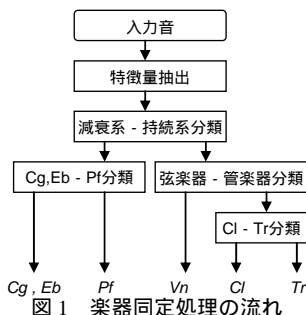


表 1 楽器の階層分類

減衰系	Classic guitar (Cg)	Electric bass (Eb)	Piano forte (Pf)
持続系	弦楽器	Violin(Vn)	管楽器
		木管	Clarinet(Cl)
		金管	Trumpet(Tr)

### 2.4 音高推定処理

ケプストラム分析で得られた周波数微細構造の情報から音高推定を行う。楽器の音高は 12 平均律で計算されるが、ケプストラム分析で計算された音高は必ずしも 12 平均律の値と一致しない。そこで以下のような補正処理をかける。観測された周波数値を  $f'$  とすると、 $f_{n-1} \times 2^{1/24} \leq f' < f_n \times 2^{1/24}$  のとき  $f' = f_n$  とする。

### 3. 評価実験

提案手法の評価実験を行った。各楽器の音高正解率を図 2 に、楽器正解率を図 3 にまとめる。正解率の平均は音高推定 87.70[%]、楽器同定がケプストラム 83.35[%]、MFCC90.98[%]であった。また、楽器毎のモデルを用いる方法と今回提案した階層的分類モデルを用いる方法との楽器同定の評価比較を図 4 に示す。この結果より、どちらの特徴量においても、階層的分類モデルを用いた場合の方が高い精度を得られていることが確認できる。

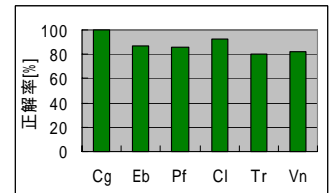


図 2 音高正解率

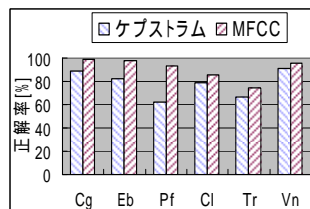


図 3 楽器正解率

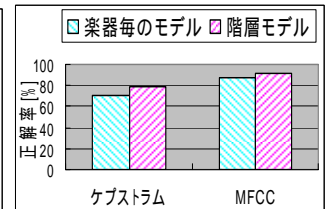


図 4 手法の比較

表 2 実験データ内訳

楽器	Cg	Eb	Pf	Cl	Tr	Vn
音域	E2-D5	E1-E3	E1-E7	D3-D6	B3-B5	G3-G6
学習データ	1296	473	655	185	143	370
テストデータ	652	224	245	106	123	142

### 4. まとめ

本稿では楽器の階層的分類を考慮した SVM により音源同定を実現する手法を提案した。その結果、楽器毎のモデルを用意する手法で認識不可能だったデータの認識が可能になり、精度向上を図ることができた。今後は音源分離を加えて和音入力への応用を行いたいと考えている。

#### 参考文献

- [1]. J. Eggink and G. J. Brown, "Instrument recognition in accompanied sonatas and concertos," Proc. ICASSP 2004.
- [2]. 北原聡志 甲藤二郎, "Support Vector Machine を用いた楽器同定における特徴量の比較," 信学会 春季全国大会, 2005.