

個体差を考慮した楽器の音源同定に関する研究*

川原三嗣 (学籍番号 200721525)

研究指導教員：田中和世

副研究指導教員：三河正彦

1. はじめに

人間の音響情報認識能力を計算機上で実現する聴覚情景解析の応用として音楽情景解析の研究がおこなわれている。楽器の音源同定(以下、楽器音同定)はその要素技術で音楽音響信号から楽器に関する情報を抽出し楽器を識別する処理である。本研究では、単楽器の孤立発音を研究の対象とした。

これまで、この研究分野[1][2]では楽器の個体差は問題視されてこなかった。個体差とは、同じ楽器における原材料、製作者、製作時期、製作場所、演奏者などその楽器を構成する要素の違いを指す。個体差により音色は変化し、同じ楽器でも識別が困難になる。そこで、本研究では楽器音同定の個体差による識別精度の劣化を改善する処理を提案し、評価実験を通じて有効性や課題、また、楽器音同定における個体差の問題について考察する。

2. 楽器音の個体差について

楽器は音楽の性質上以下の特徴がある。

- 所望の音程、音色を得るために種類によって異なった構造、大きさを持っている
- 同じ楽器種であっても各部の細かな構造、構成材料が異なる場合がある
- 材料として木材や皮のような均一性の確保が難しい素材が使われる

以上の理由から、同じ楽器種でも個体差が生じ、その影響で音色も変化する。また、影響は楽器種によって異なることが予想される。そのため、従来と同様、識別処理を一括して行った場合、ある楽器には有効であるが他の楽器には有効でない特徴量が含まれる可能性がある。

3. 提案手法

そこで、本研究では識別処理を分割し、各処理部で適切な特徴量を選択することを考える。そのための特徴量として、サブバンド出力特徴を、識別処理として階層的識別手法と各階層での特徴量選択処理を提案する。

3.1 サブバンド出力特徴

サブバンド出力特徴は音響信号のパワースペクトルをサブバンドに分割し、それぞれのバンドのパワーの時間変化を抽出したものである[3]。ケプストラムに似た特徴量であるが、特徴ベクトルの要素と周波数との対応が明瞭で選択処理に有効である。なお、本稿で述べるサブバンドは、ガウス型のフィルタ群を利用して最大バンドの中心周波数が 8[kHz] でバンド数が 128、すなわち中心周波数間隔が 66[Hz] となるサブバンドを使用した。時間変化についてはおよそ 250[ms] 分を用いた。

3.2 階層的識別手法と特徴量選択処理

階層的識別手法はまず、楽器の発音機構・形状などから階層を設定し、その上位から下位へ順次識別を行い、最終的に到達した楽器を識別結果とする[4]。なお、各階層で用いる識別器は k -NN 法とした。

各階層で有効なバンドを自動選択する手法として重み付きサブバンド出力特徴を用いた主成分分析(PCA)を採用する[5]。PCAを行うデータはサブバンド出力特徴の各バンドの時間平均である。累積寄与率が 0.9 以上となる主成分までを使用し、それらの主成分係数を用いてサブバンド圧縮を行うことで近似的にバンド選択を実現する。PCAでは分散が大きい次元を重要な次元と仮定する。そのため低音域の分散が極端に大きい楽器音デ

* “Studies on Musical Instrument Identification Focusing on Individuality of Instruments” by Mitsugu KAWAHARA

ータでは、分散が低い重要度は高いバンドが小さい重みになる可能性がある。そこで本研究では PCA の前処理としてサブバンド出力特徴に重み付けを行う。重み関数については予備実験で最もよい結果を得たものを採用した。

4. 評価実験

4.1 実験概要

楽器音データベース[6]の 18 の楽器種を用いた評価実験を行った。本稿では以下の 4 つの実験の結果を示す。

A) 個体差を考慮しないデータセットに対して非線形 SVM(Support Vector Machine) で一括識別を行った場合

B) 個体差を考慮したデータセットに対して非線形 SVM で一括識別を行った場合

C1) B と同じデータセットに対して 3-2 で説明した階層的識別手法を用いた場合

C2) B, C1 と異なる個体差を考慮したデータセットを用いて C1 と同手法を用いた場合

なお、特徴量はすべて 3-1 で述べたサブバンド出力特徴である。

4.2 結果と考察

結果を Table 1 に示す。総合はすべてのサンプルデータのうちの正解数で、平均は楽器毎の識別率の平均を表す。まず、A の結果よりサブバンド出力特徴が楽器音同定で従来研究と同等以上の識別精度を持つことがわかった。しかし、同じ識別手法でも個体差を考慮した B の場合、識別精度が大幅に劣化する。これに対して階層的識別手法を用いた C1 では、全体の個体差による識別精度の劣化が改善されている。ただし、楽器種によっては B の方が良いものも含まれる。また、個体を変えた C2 の結果は C1 とは大きく異なる。これより個体差の影響が複雑で、少数の個体差から学習できる傾向が他の個体には当てはまらない場合が多いことが予想される。

5. 結論

個体差を考慮した楽器音同定について述べた。今後の課題として、識別精度のさらなる改善

Table 1 実験結果

楽器	A	B	C1	C2
PF	100.0	99.6	92.9	57.1
CG	98.6	12.9	14.8	97.1
AG	98.3	79.5	69.0	79.5
EG	99.8	98.6	87.1	75.7
VN	98.3	28.3	56.1	97.2
VL	96.1	86.1	87.8	55.0
VC	99.1	93.9	84.4	83.9
TR	96.6	55.2	36.2	10.0
TB	100.0	100.0	90.0	100.0
SS	97.0	38.9	61.1	67.8
AS	97.0	0.0	52.2	24.4
TS	96.3	2.2	30.0	50.0
BA	99.3	0.0	24.4	57.8
OB	95.8	43.0	60.0	51.1
FG	98.5	94.3	83.8	76.2
CL	95.8	97.5	87.5	84.2
PC	98.9	93.3	98.9	94.4
FL	95.7	12.4	17.1	49.5
総合	98.2	62.6	65.3	70.2
平均	97.9	57.5	63.0	67.3

のために各階層での識別処理の最適化を行う必要がある。また、楽器音データベースの充実も含めた、楽器の個体差に関するさらなる研究が必要である。

参考文献

- [1] 北原鉄朗他, 情報処理学会論文誌, D-14-28, 2448-2458, 2003.
- [2] 馬場貴之他, 情報処理学会研究報告, 2005-MUS-61, 79-84, 2005.
- [3] 川原三嗣他, 音響学会講演論文集(春), 2-5-14, 917-918, 2008-3.
- [4] 川原三嗣他, 音響学会講演論文集(秋), 1-9-23, 905-906, 2008-9.
- [5] 川原三嗣他, 音響学会講演論文集(2009.3 発表予定)
- [6] 後藤真孝他, 情報処理学会研究報告, 2002-MUS-45, 19-26, 2002.