

# くし形フィルタと自己相関関数による演奏楽器推定\*

北見伸一郎 東海林智也†

函館工業高等専門学校 情報工学科‡

## 1 はじめに

入力された楽曲の演奏音声から楽譜を作成する自動採譜システムでは同時に演奏楽器推定もおこなう必要がある。従来の楽器推定手法では楽器音のスペクトルをテンプレートとして使用することが多いが、その場合は楽器数が増えるにつれてテンプレート数とファイルサイズも増えるという問題がある。そこで本研究では、くし形フィルタを用いることによりテンプレート数を減らし、さらに自己相関関数を用いてテンプレートマッチングをおこなうことでファイルサイズを減らす演奏楽器推定システムの構築を検討する。

## 2 くし形フィルタ

整数  $i$  をオクターブ、整数  $p$  を音名番号 ( $p = 1, 2, \dots, 12$ , 1 が C, 12 が B に対応) とすると、任意の音は  $i$  と  $p$  の組  $(i, p)$  で表される。例えば (3, 1) は O3C, (4, 2) は O4C# に対応する。また  $f_{i,p}$  (Hz) を  $(i, p)$  音の基本周波数とし、 $f_s$  (Hz) をサンプリング周波数とする。このとき  $(i, p)$  音に対応する正整数の定数として  $N_{i,p} = [f_s/f_{i,p}]$  ( $[]$ : 整数への端数処理) を定義すると、 $(i, p)$  音に対応する Resonator 型くし形フィルタは次の伝達関数で表される [1]

$$H_{i,p}(z) = \frac{1-a}{1-a \cdot z^{-N_{i,p}}} \quad (1)$$

このフィルタは出力  $y_{i,p}(n)$  を  $N_{i,p}$  時刻だけ遅延させて入力に加算するだけの単純な IIR 型フィルタであり、利得特性  $|H_{i,p}(z)|$  は図 1 で表される。図 1 から、Resonator 型くし形フィルタは  $(i, p)$  音に含まれる基本周波数とその倍音成分の周辺を残し、その他の成分を減衰させる特性を持つことが分かる。なおフィルタ係数  $0 \leq a < 1$  は任意のパラメータであり、 $a$  が 0 に近づくほど利得は平坦になり、1 に近づくほど急になる。

この特性より  $(i, p)$  音に対応するくし形フィルタ  $H_{i,p}(z)$  に任意の和音を通して得た出力  $y_{i,p}(n)$  には  $(i, p)$  音の成分が多く含まれている。よって、和音を分離しないで全ての楽器音の組み合わせに対してテンプレートマッチングをおこなうよりも、くし形フィルタからの出力に対して単一楽器に対してテンプレートマッチングをおこなうことでテンプレート数を減らすことが出来る。

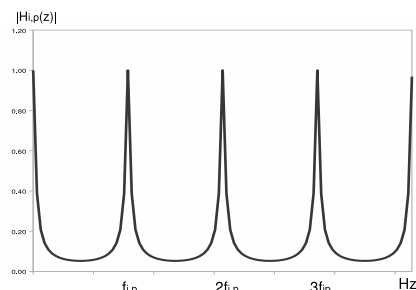


図 1: Resonator 型くし形フィルタの利得 ( $a = 0.9$ )

なお、楽器推定ではなく音高推定を主な目的とする場合は  $(i, p)$  音の成分を残す特性を持つ Resonator 型よりも、 $(i, p)$  音の成分を除去する特性を持つ Notch 型くし形フィルタを用いることが多いが [2]、本研究ではフィルタ出力に対してテンプレートマッチングをおこなうため Resonator 型を採用した。

## 3 演奏楽器推定システムの構成

本研究で提案する演奏楽器推定システムは次の手順で動作する。ここで時刻  $n$  におけるシステムへの入力音声信号を  $x(n)$  とし、 $(i, p)$  音に対応するくし形フィルタを  $H_{i,p}(z)$  とする。

- (1) オクターブ  $i$  の範囲を決定し、全ての  $(i, p)$  の組に対応した  $H_{i,p}(z)$  を作成してそれらを並列接続する (図 2)。時刻  $n$  ごとに  $x(n)$  を並列接続された全ての  $H_{i,p}(z)$  に入力し、それぞれのフィルタ出力  $y_{i,p}(n)$  を得る。
- (2) ある時間間隔ごとに、全ての  $y_{i,p}(n)$  に対する短時間での標準偏差  $s_{i,p}$  を求める。
- (3) 今回は音高推定が目的ではないので、オクターブの違いを無視することにして各  $p$  音に対して標準偏差の和  $S_p = \sum_i s_{i,p}$ ,  $p = 1, 2, \dots, 12$  を求める。ここで、ある  $S_p$  が任意に定めたしきい値  $TH_S$  よりも大きい場合は、あるオクターブにおいて  $p$  音が演奏されていると判断することが出来る。
- (4) 演奏されていると判断した  $p$  音に対して、各オクターブ別に  $y_{i,p}(n)$  から自己相関関数  $r_{i,p}(k)$ ,  $k = 1, 2, \dots, R_{\max}$  を求めてテンプレートマッチングをおこなう。ここで  $R_{\max}$  は自己相関関数長である。今回はテンプレートマッチングとして  $r_{i,p}(k)$  とテンプレートとの相関を使用するので、マッチングの指標を  $R_{i,p}(l)$  ( $l$ : 楽器番号) とおくと、 $R_{i,p}(l)$

\*A Musical Instrument Estimation Using both Comb Filters and Auto Correlation Functions

†tokai@hakodate-ct.ac.jp

‡Department of Computer Engineering, Hakodate National College of Technology

の最大値は 1, 最小値は  $-1$  となる. さらに  $l' = \operatorname{argmax}_l R_{i,p}(l)$  とすると, もし  $R_{i,p}(l')$  が任意に定めたしきい値  $\text{TH}_R$  より大きい場合は  $p$  音は楽器  $l'$  で演奏されていると判定出来るので楽器番号  $l'$  を記録する.

- (5) (3) と (4) を  $p = 1, 2, \dots, 12$  に対して繰り返す.
- (6) 短時間で楽器が頻繁に切り替わることはあまり考えられないので, 楽曲のフレーズ単位で楽器認識をおこなう. フレーズ終了時に記録した楽器番号の頻度を計算し, 頻度が大きい楽器の順に楽器名を示す.

なお (4) において FFT ではなく自己相関関数を用いてテンプレートマッチングをおこなっている理由は, 特に中低音域において十分な周波数分解能を得るために, 自己相関関数よりも FFT の方が長時間の音声信号が必要だからである [3].

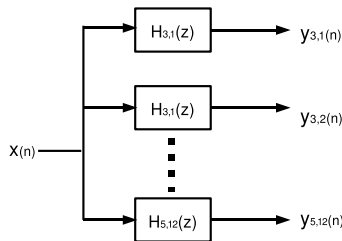


図 2: くし形フィルタの並列接続 ( $i = 3, 4, 5$ )

## 4 テンプレートの作成

提案する演奏楽器推定システムでは, 入力音声信号はくし形フィルタに通されて利得の影響を受けるため, 通常の楽器入力音から自己相関関数を求めてテンプレートとして用いることは出来ない. そこで楽器の  $(i, p)$  音をくし形フィルタ  $H_{i,p}(z)$  に通して求めた自己相関関数を  $(i, p)$  音に対するテンプレートとして用いることにする (図 3).

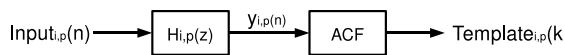


図 3: テンプレートの作成方法

## 5 実験

### 5.1 実験条件

電子音 (MIDI) を用いて提案する演奏楽器推定システムの実験をおこなった. 入力音声信号は MIDI 出力を変換して作成した wave ファイル (44.1Hz, 16bit, モノラル) とした. 対象楽器はピアノ (PF), トランペット (TR), アルトサクソ (AS), バイオリン (VN), フルート (FL), エレキベース (EB) の 6 種類とし, オクターブ範囲  $i = 3, 4, 5$  で定常状態にある音を用いてテンプレートを作成した. また, テンプレートマッ

グをおこなうタイミングは 50 ミリ秒おきとし, しきい値を  $\text{TH}_S = 2000$ ,  $\text{TH}_R = 0.8$  と設定した. 入力音声信号は 2 種類の楽器 (PF, TR) からなるワンフレーズ音声 (約 8 秒) とした.

### 5.2 結果と考察

くし形フィルタの係数  $a$  の値を変えてフレーズ内での楽器の判定頻度を調べた. 実験結果を表 1 に示す.

a	PF	TR	AS	VN	FL	EB
0.90	37.8	34.4	4.5	7.3	7.7	8.3
0.95	39.5	38.0	1.9	4.6	9.2	6.8
0.99	23.6	32.6	1.3	2.0	26.3	13.9

表 1: 各楽器の頻度 (単位:%)

以上の結果から  $a = 0.95$  の場合にもっとも PF と TR の頻度が高くなった. また  $a = 0.90$  の場合は PF と TR の頻度は下がった. これはくし形フィルタの利得特性が平坦になり過ぎてうまく和音分離がされなかったからと考えられる. 逆に  $a = 0.99$  の場合にも PF と TR の頻度は下がった. これはくし形フィルタの利得特性が急すぎて楽器音の判定に必要な音声成分まで取り除いてしまったからと考えられる. したがって楽器推定をうまくおこなうためには最適な  $a$  を定める必要があることが分かった.

## 6 今後の課題

今回の実験では経験的にフィルタ係数やしきい値を決定したが, より正確に楽器推定をおこなうためには最適な値を自動的に見つける手法を検討する必要がある. また, 電子音ではなく実際の楽器音を用いた実験, 演奏楽器数を増やした実験, 演奏楽器の組み合わせを変えた実験をおこなう必要がある.

謝辞 本研究は, 豊橋技術大学高専教育連携プロジェクト「採譜システムにおける楽器推定法に関する研究」の援助によりおこなわれた.

### 参考文献

- [1] 藤原道, 山口満, 斎藤努, 田所嘉昭, Resonator 型くし形フィルタによる打楽器音を含む楽器音の音高推定法の検討, 電子情報通信学会技術研究報告, pp.19-23, EA2003-80, 2003.
- [2] 坂内秀幸, 夏井雅典, 田所嘉昭, くし形フィルタに基づく自動採譜システムの実現, 情報処理学会研究報告, pp.13-18, 2007-MUS-71, 2007.
- [3] 瀬戸康裕, 夏井雅典, 田所嘉昭, くし形フィルタと相関関数による音高推定困難和音の音高推定法, 情報処理学会研究報告, pp.1-6, 2006-MUS-68, 2006.