

## F-18

## 並列構成くし形フィルタの最小出力による楽音の音高推定の検討

## Consideration of Pitch Estimation for Music Sounds Noticing

## Minimum Output of Comb Filters Connected in Parallel

田所 嘉昭†  
Yoshiaki Tadokoro森田 健夫†  
Takeo Morita

## 1. まえがき

採譜において音高推定は重要な役割を果たす[1]-[3]. 楽音は、基本周波数とその倍音から構成され、それらの成分は、音高、楽器等によって複雑に変化し、そのことが音高推定を困難にしている. 従来の音高推定法は、主に高速フーリエ変換や適応フィルタにより基本周波数を抽出するものであるが、楽音の性質のため、かなり複雑な処理を必要としている[4]-[7]. 著者らは、入力サンプル値の減算で実現できるくし形フィルタ ( $H(z)=1-z^{-N}$ ) を基本にした基本周波数と倍音成分を消去する音高推定法を提案し、異種複数楽器による多重音の音高推定が可能であることを示した[8],[9]. この原理は、くし形フィルタを1オクターブの音名数(12個)だけ縦続接続し、その零出力を検出することにより行われる. しかし、楽音周波数の変動や雑音成分が含まれると縦続くし形フィルタの特性より、零出力検出が困難になる問題点がある. これを改善する方法として、くし形フィルタを12個並列に接続し、どのフィルタ出力において周波数成分が減少しているかを検出することで音高を推定する方法を提案した[10]. これはくし形フィルタの各出力を特異値分解(SVD)して、その周波数成分を検出することにより行われる. しかし、SVDの計算量は入力データ( $N \times N$ )のとき、 $N^3$ の乗算数を必要とする.

本論文では、並列くし形フィルタで1音が除かれたくし形フィルタの出力値は他のフィルタ出力に比較して振幅が小さくなるという仮定の下に、音高推定アルゴリズムを検討した. これにより演算量の削減が期待できる.

## 2. 音高推定法

オクターブ  $p$ , 音名  $q$  (ピッチ周波数  $f_{pq}$ ) のくし形フィルタは式(1)で示され、Fig.1に示されるような周波数特性を持つ. これより、 $f_{pq}$ とその倍音成分を除去可能である.

$$H_{p,q}(z) = 1 - z^{-N(p,q)} \quad (1)$$

ただし、サンプリング周波数を  $f_s$  とするとき、 $N(p,q)$  は次式で決定される.

$$N(p,q) = [f_s / f_{pq}] \quad [ ] : \text{整数化} \quad (2)$$

採譜対象音域をオクターブ3から5としたとき、オクターブ3の12個の音名に対応するくし形フィルタをFig.2(a)のように並列接続する. そして、各出力の一定期間での最大出力値  $|y_{q,\max}|$  を求める. そして、この  $|y_{q,\max}|$  ( $q=1,2,\dots,12$ ) の最小値に対応するくし形フィルタ  $H_{3,m1}(z)$  を検出し、Fig.2(b)のようなフィルタ構成にする. (b)においても上述と同じ処理を行い、最小出力を示す  $H_{3,m2}(z)$  を検出し、 $H_{3,m1}(z)$  の前に接続する. この処理を並列くし形フィルタのどれかの出力が零出力になるまで行うことで、入力音の音名が推定される.

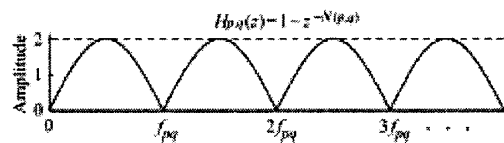
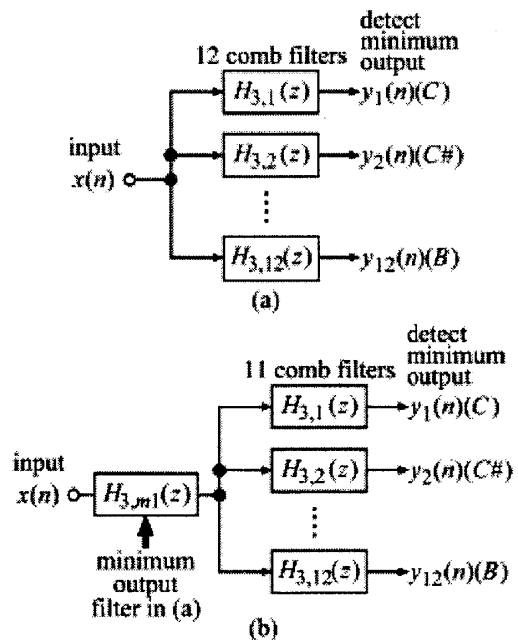
Fig.1 くし形フィルタ  $H_{p,q}(z) = 1 - z^{-N(p,q)}$  の振幅特性

Fig.2 並列構成くし形フィルタによるピッチ検出システム

各音名に対するオクターブはくし形フィルタの次数  $1/2$  にていくことで ( $N(3,q)/2 = N(4,q)$ ,  $N(4,q)/2 = N(5,q)$ ) 推定できる[8]. これより音高が推定される.

## 3. 疑似楽音による音高推定特性

提案法の特性を調べるため、 $r$  音の各音に対し、6倍音まで考慮した式(3)に示す疑似楽音に対する音高推定特性を調べた.

$$x(n) = \sum_{j=1}^r \sum_{k=1}^6 A_{jk} \cos((2\pi k f_j n / f_s) + \phi_{jk}) \quad (3)$$

ただし、 $f_j$  は  $j$  音の基本周波数である.

オクターブ3内の2音の組合せ  $66(12C_2)$ , 3音の組合せ  $220(12C_3)$  およびオクターブ3から5内の2音の組合せ  $630(36C_2)$ , 3音の組合せ  $7,140(36C_3)$  の多重音に対して、 $A_{jk}, \phi_{jk}$  を  $0 \leq A_{jk} \leq 1.0, 0 \leq \phi_{jk} \leq 2\pi$  の範囲でランダムに変化

†豊橋技術科学大学

させて  $10^4$  個の入力音を作成した。これらの入力音の 100 ms 間の並列くし形フィルタ出力の最大値  $|y_{q,\max}|$  を検出し、2 節のアルゴリズムに従って、音高を推定した。その推定結果を Table 1 に示す。オクターブ 3 に限れば、2 音で 98.3%、3 音で 91.1% の確率で検出できている。オクターブ 3 から 5 における多重音に対しては、オクターブ違い、協和音を含む重複関係音の区別が完全にできていないため、

Table 1 ピッチ推定結果

input	octave 3	octave 3 to 5
one pitch	100%	100%
two pitches	98.3%	90.6%(93.3%)
three pitches	91.1%	72.0%(86.8%)

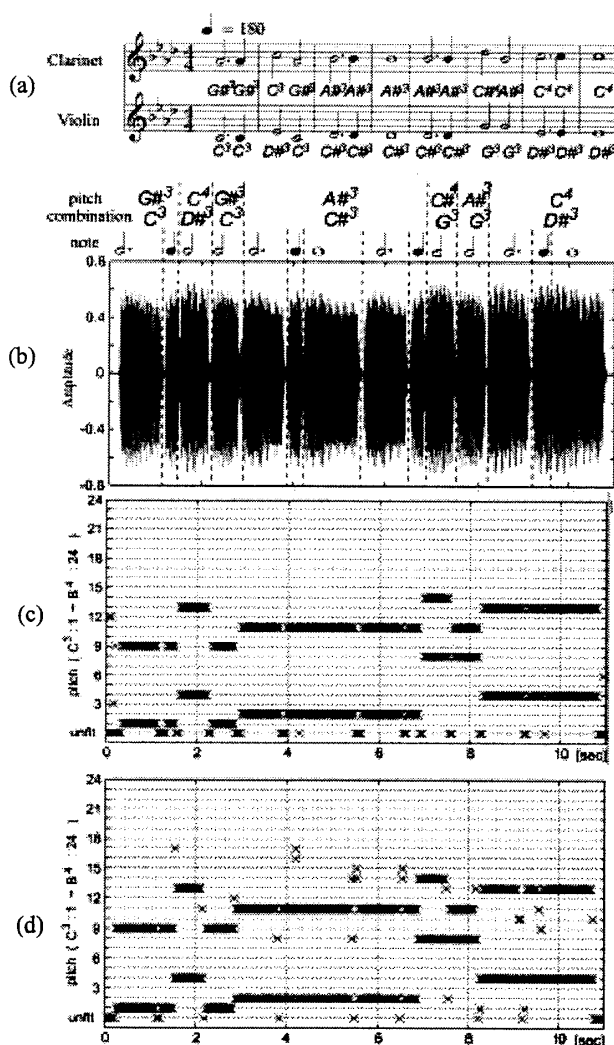


Fig.3 (a) テスト楽譜(It's small world), (b) 入力波形, (c) 提案法による音高推定結果, (d) SVD による音高推定結果

確率が 2 音で 90.6%, 3 音で 72.0% と低くなっている。

( ) 内の数値は、ある音のスペクトル成分のすべてが他の音のスペクトルに含まれるような和音を除いたときの確率である。

#### 4. 電子楽器音に対する音高検出結果

図 3(a) に示す楽曲をクラリネット、バイオリンで演奏した場合の楽曲に対する音高推定結果を同図(c) に示す。ほぼ正しい結果が得られており、(d) に示す SVD を用いた結果とあまり変わらない。また、実際に良く使われる長三和音、短三和音の 34 個の和音については、95.5% の検出率を示した。

#### 5. むすび

並列構成くし形フィルタの各出力の一定期間中の最大値を求め、その中の最小値を示すくし形フィルタを選択していくことによって音高を推定する方法を提案した。オクターブ 3 内の 2 音、3 音に対して、それぞれ 98.3%、91.1% で音高推定ができることを示した。オクターブ 3 から 5 の 2 音で 90.6%、3 音で 72.0% となった。また、楽曲を用いたシミュレーションにより、SVD による方法とほぼ同程度の結果が得られることを確認した。提案法では、出力値の大小関係の処理に対し、SVD では  $O(N^3)$  の乗算を必要とする。

今後は、検出不可能な和音の解決法と、周波数変動に対する特性を検討する。

#### 参考文献

- [1] M. Piszczalski and B. A. Galler, "Automatic music transcription," Computer Music Journal, vol. 1, no. 4, pp. 24-31, 1977.
- [2] C. Roads, "Research in music and artificial intelligence," ACM computing Surveys, vol. 17, no. 2 pp. 163-190, 1985.
- [3] 片寄, 井口, "知的採譜システム," 人工知能学会誌, vol.5, no.1, pp.59-66, 1990.
- [4] J. A. Moorer, "Signal processing aspects of computer music: a survey," Proceedings of IEEE, pp. 1108-1137, 1977.
- [5] W. J. Pielemeier and G. H. Wakefield, "A high-resolution time-frequency representation of musical instrument signals," J. Acoust. Soc. Am. vol. 99, no. 4, Pt. 1, pp. 2382-2396, 1996.
- [6] S. Foster, J. Rockmore, W. Schloss, "Toward an intelligent editor of digital audio: signal processing methods," Computer Music Journal, vol. 6, no. 1, pp. 142-61, 1982.
- [7] 梅本, 青島, "適応技術を用いた採譜に有効な周波数分析法," vol.28, no.5, pp.619-626, 1992.
- [8] 三輪, 田所, 斎藤, "くし形フィルタを利用した採譜のための異楽器音中のピッチ推定," 信学論, vol.J81-D-II, no.9, pp.1965-1974, 1998.
- [9] Y. Tadokoro and T. Miwa, "Musical pitch and instrument estimation of polyphony using comb filters for transcription," CSCC200, Advances in Physics, Electronics and Signal Processing Applications, pp.315-319, 2000.
- [10] Y. Tadokoro, W. Matsumoto and M. Yamaguchi, "Pitch detection of musical sounds using adaptive comb filters controlled by time delay," IEEE International Conference on Multimedia and Expo, (to be accepted), Lausanne, Switzerland, 2002.
- [11] M. Yamaguchi and Y. Tadokoro, "Pitch detection of song sounds using parallel connected comb filters and singular value decomposition," IEEE International Symposium on Intelligent signal processing and communication systems (ISPA CS), (to be submitted), Taiwan, 2002.