

## 7段縦続接続くし形フィルタによるピアノ和音の音高推定法の検討

松山 大仁郎, 夏井 雅典, 田所 嘉昭  
(豊橋技術科学大学)

自動採譜の実現のために、縦続接続したくし形フィルタを用いた多和音の音高推定法について検討を行う。1 オクターブ 12 音に対応する 12 段縦続接続くし形フィルタを用いた音高推定法では、縦続数の増加により、接続の後段のくし形フィルタに対応する音高の推定が困難であるという問題が生じる。そこで、音高の倍音関係を利用し、縦続接続数を減少させた音高推定法として、7 段縦続接続くし形フィルタを用いた音高推定法を提案する。さらに、後段のくし形フィルタに対応する音高の推定を、high-pass filter (HPF) と仮推定という手法で補正を行う。実際にピアノ 4 和音を推定したところ、約 77% の推定率を得た。

### Consideration of Pitch Estimation Method for Piano Chords Consisting of Many Notes Using Cascaded Seven Comb Filters

Daijiro Matsuyama, Masanori Natsui, Yoshiaki Tadokoro  
(Toyohashi University of Technology)

To realize the music transcription using a computer, we propose a pitch estimation method for chords consisting of many notes using cascaded comb filters. In a pitch estimation method using twelve cascaded comb filters, it is difficult to estimate the pitches corresponding to the latter comb filters because of the frequency characteristic of the cascaded comb filters. We use the cascaded seven comb filters and a high-pass filter (HPF) and a temporary estimation to estimate piano four chords. As a result, we obtained about 77% estimation accuracy.

#### 1. ま え が き

音楽信号を楽譜に書き起こす処理のことを採譜と呼び、この作業を計算機によって自動的に行う技術が自動採譜である。ここで、自動採譜システムを計算機上に実現するためには、楽音の音高を推定する処理が必要となる。音楽信号は、基本周波数成分とその整数倍の周波数を有する多数の高調波成分から構成されており、音高推定とは基本周波数を推定する処理のことを指す<sup>1)~4)</sup>。

自動採譜の実現を目的として、本研究室ではくし形フィルタ ( $H(z) = 1 - z^{-1}$ ) を利用した音高推定法を検討している<sup>5)~7)</sup>。くし形フィルタは、周波数軸上に等間隔の零点を有しており、単音による音楽信号に含まれる全調波成分を除去する性質を有する。この性質を利用し、各音高に対応するくし形フィルタの出力から零出力を検出することにより、音高の推定が可能となる。また、くし形フィルタは遅延器のみを用いて構成されるため、時間領域においてはサンプル点間の差分のみで処理が可能であり、演算量の削減および処理系の単純化を図ることができる。

本研究室では、これまで、1 オクターブ 12 音に対応するくし形フィルタを並列に接続し、その各フィルタの出力の振幅値に注目した音高推定法を提案している<sup>7)</sup>。本

手法を用いることで、ピアノ 2 和音は約 95%、3 和音では約 80% の推定率で音高を推定できることを確認している。しかし、和音の構成音数が増えるにつれ、各フィルタの振幅値の比較が難しくなり、入力信号が 4 和音以上になると、推定率が約 60% まで下がることを確認している。

次に、くし形フィルタを 12 段縦続接続したシステムを用い、各フィルタの出力の振幅値と時間波形に注目した音高推定法を提案している<sup>8)</sup>。この手法では、原理的に、全ての音高を消去した時点で零出力を検出するので、和音の構成音数が増えても推定率は変わらないと考える。しかし、実楽器音では、くし形フィルタの縦続接続数が多くなるほど、そのフィルタの振幅特性により低域の雑音が増幅し、音高の推定が困難となることを確認している。

そこで、本稿では、くし形フィルタの縦続接続数を減らした音高推定法として、7 段縦続くし形フィルタを用いた手法について検討を行う。音高の倍音関係を利用し、くし形フィルタの縦続接続数を減らすことで、音高推定時に問題となる低域の雑音成分の増加を抑制できる。なお、本論文では、オクターブ 4 の音域 (C4: 261.62Hz ~ B4: 493.88Hz) におけるピアノ音 (RWC の音楽データベース<sup>9)</sup>) に限定して検討を行うが、この音名処理にオクターブ推定の処理<sup>7)</sup>を追加させると、推定音域は拡大できる。

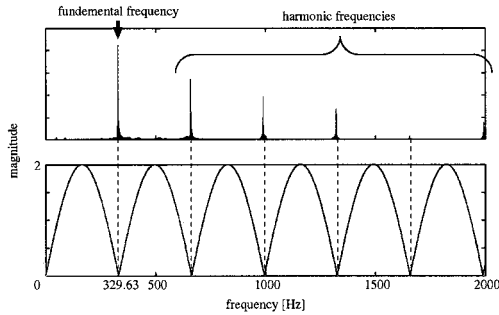


図1 ピアノ E4 の周波数特性とくし形フィルタの振幅特性

## 2. くし形フィルタによる音高推定

### 2.1 くし形フィルタによる音高推定の基本原理

オクターブ  $i$ , 音高  $p$  の基本周波数を  $f_{i,p}$  とした場合, 音楽信号は基本周波数成分と, その整数倍の高調波成分から構成される. ここで,  $p = 1(C), 2(C\#), \dots, 12(B)$  である. この音楽信号を消去するくし形フィルタの伝達関数  $H_{i,p}(z)$  は式 (1) で表される. また, このフィルタに信号  $x(n)$  を入力した時の出力  $y_{i,p}(n)$  は式 (2) で表される.

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

$$y_{i,p}(n) = x(n) - x(n - N_{i,p}) \quad (2)$$

ここで,  $N_{i,p}$  は遅延数であり, サンプリング周波数を  $f_s$  とすると  $N_{i,p}$  は式 (3) で求められる.

$$N_{i,p} = \left\lceil \frac{f_s}{f_{i,p}} \right\rceil \quad [ \cdot ] : \text{integer by rounding} \quad (3)$$

例として, 図1にオクターブ4の音名 E のピアノ楽音 (E4:  $f_p = 329.63\text{Hz}$ ) の周波数特性と, 対応するくし形フィルタの振幅特性を示す.

くし形フィルタは, 周波数領域において等間隔に零点が存在するような特性を有しており, 音楽信号を構成する複数の周波数成分を同時に除去することができる. そのため, くし形フィルタの出力値から零出力を検出することで, 音高の推定が可能となる.

### 2.2 並列構成くし形フィルタ

くし形フィルタの性質を利用し, 1オクターブ12音に対応するくし形フィルタを並列に接続したシステム (並列構成くし形フィルタ) を用いて和音の推定を行う. 入力信号  $x(n)$  が単音である場合, 各くし形フィルタの出力のいずれかは零値を出力するため, その出力値から単音の推定が可能である. 入力が和音である場合, 各くし形フィルタの出力は全て非零値となるが, 各出力を比較し, それらの最小出力に注目することで, 入力中のいずれか1音を推定できる.

並列構成くし形フィルタの最小出力に注目した音高推

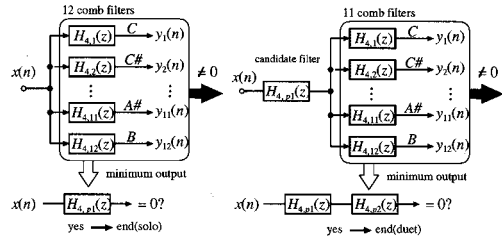


図2 PCF法

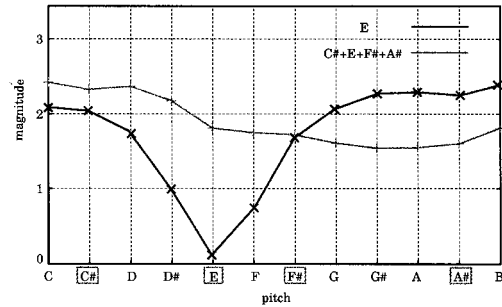


図3 和音数の増加による各出力値の変化

定法として PCF 法 (Parallel Comb Filter 法) が提案されている<sup>7)</sup>. 推定手順を図2に示す. PCF 法では, 入力単音であれば零出力から1音を推定する. 入力及和音であれば, 各フィルタの最小出力値から入力及音中の1音を推定し, 推定した音高に対応するくし形フィルタをシステムの前段に移動し, 同様に信号を入力して零出力が得られるまで音高を推定していく. 零出力の判定は, くし形フィルタの入出力パワー比  $R$  を式 (4) により求め, しきい値よりもパワー比が小さければ零出力とみなす. ここで,  $N_{all}$  はくし形フィルタの総遅延数であり,  $N_T$  は 10ms 分のデータ (441 点) としている.

$$R = \frac{\sum_{nT=0}^{N_T-1} y(n+nT+N_{all})^2}{\sum_{nT=0}^{N_T-1} x(n+nT)^2} \quad (4)$$

PCF 法を用いることで, オクターブ4のピアノ3和音の音高推定が推定率80%で可能であることを確認している. しかし, 入力信号が4和音となると, 並列構成くし形フィルタに通して1音を消去しても, 出力値には3和音残っているため, 各出力値の減衰があまりみられず, 各出力値に差が生じにくく, 1音の推定が困難となる. その結果, 4和音の推定率は約60%まで低下することを確認している. 図3に出力値の一例を示す. 単音を入力した場合は, 入力音である E の出力値が最も小さくなっているが, 4和音を入力すると, 出力値に差があまり見られなく, 入力音ではない音の出力値が最小となっている.

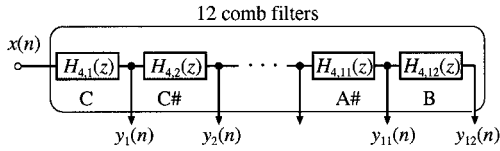


図4 12段縦続接続くし形フィルタ

### 2.3 縦続接続くし形フィルタ

並列構成くし形フィルタでは和音数が4和音以上になると推定が難しくなるため、和音数に依存しない推定法として、縦続接続くし形フィルタに注目する<sup>5)</sup>。ここで、1オクターブ12音に対応するくし形フィルタを図4のように12段縦続接続したシステムについて考える。

原理的には、入力  $x(n)$  に含まれる音を1音ずつ除去し、入力中の全ての調波成分を除去した段階で、零値を出力する。この全音消去時の零出力を検出することで、和音の音高推定が可能になる。図5に12段縦続接続くし形フィルタの処理の流れを示し、以下に説明を行う。

- 1) 縦続接続くし形フィルタに信号を入力。
- 2) 各くし形フィルタの出力のパワー比を式(4)より求める。求めたパワー比を比較し、パワー比が最小となるフィルタを検出。
- 3) 検出したフィルタを最前段へ移動。
- 4) 信号を入力し、同様にパワー比が最小となるフィルタを検出。すでに移動したフィルタを検出した場合は5)へ、そうでなければ3)へ戻る。
- 5) 前段へ移動させたフィルタに対応する音高を入力和音の音高とする。

縦続接続くし形フィルタの最終出力の振幅特性を図6に示す。実際の楽音では非定常性、周波数変動のため、完全な零出力を得ることは難しい。そのため、図6の振幅特性によって、縦続接続数が増えるほど低域の周波数成分が増幅され、後段になるほど信号成分の有無の確認が難しく、推定が困難になるという問題が生じる。

## 3. 提案する音高推定法

### 3.1 7段縦続接続くし形フィルタ

縦続接続数の増加が問題となるため、縦続接続数を減らしたシステムについて検討を行う。音高の関係を考えると、平均律音階では、隣接する音高間の周波数比が  $2^{1/12}$  倍と不等間隔であるため、以下のような周波数関係が成り立つ。

$$2 \times f_{3,p} = f_{4,p} \quad (5)$$

$$3 \times f_{3,p} = f_{4,p+7} \quad (6)$$

そのため、図7のように、くし形フィルタ  $H_{3,p}(z)$  は、オクターブ4の  $p$  音と  $p+7$  音を消去する。この関係を利用すると、表1に示す倍音関係から、オクターブ4の12音をC3, C#3, D3, D#3, E3, F4, F#4の7個のく

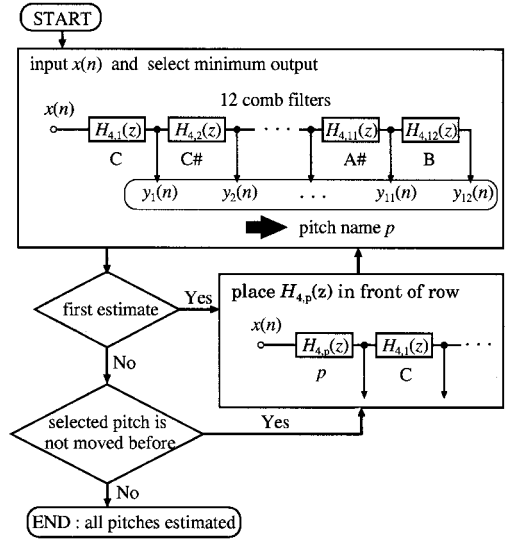


図5 縦続接続くし形フィルタによる音高推定

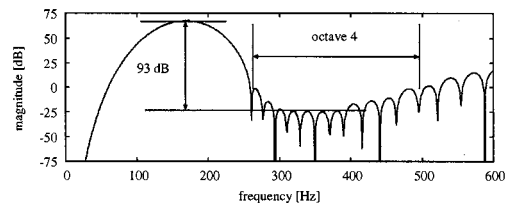


図6 12段縦続接続くし形フィルタの振幅特性

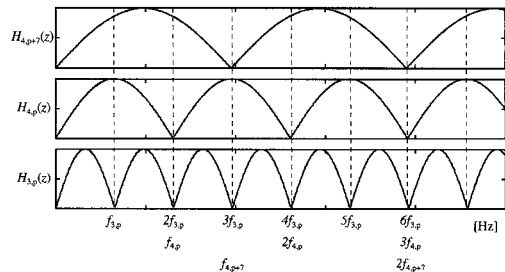


図7 くし形フィルタ  $H_{3,p}(z)$ ,  $H_{4,p}(z)$ ,  $H_{4,p+7}(z)$  の零点

し形フィルタを用いて消去することが可能である。本稿では、図8のように、くし形フィルタを7段縦続に接続したシステムを用いて和音の音高推定を行う。ここで、7段縦続接続した時の振幅特性を図9に示す。

12段縦続接続くし形フィルタと比べて縦続接続数が少ないため、振幅特性によって低域の雑音成分が増加する前に、振幅値に注目した手法での音高推定が可能であると考えられる。7段縦続接続くし形フィルタを用いた音高推定手順は、次節で示す。

表 1 音高の倍音関係 [Hz]

		2 倍音		3 倍音	
C3	: 130.81	C4	: 261.62	G4	: 391.99
C#3	: 138.59	C#4	: 277.18	G#4	: 415.30
D3	: 146.83	D4	: 293.66	A4	: 440.00
D#3	: 155.56	D#4	: 311.13	A#4	: 466.16
E3	: 164.81	E4	: 329.63	B4	: 493.88

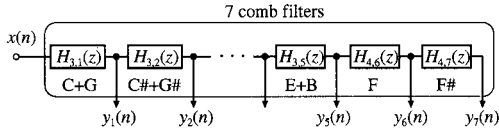


図 8 7 段縦続接続くし形フィルタ

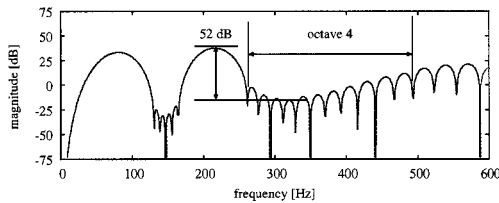


図 9 7 段縦続接続くし形フィルタの振幅特性

### 3.2 音高推定アルゴリズム

7 段縦続接続くし形フィルタを用いた和音の音高推定アルゴリズムを以下に示す。音高の推定手順は、原理的に図 5 の推定手順と同じである。ただし、1 個のくし形フィルタが 2 個の音高に対応する部分の処理が異なる。

- step 1. 7 段縦続接続くし形フィルタに信号を入力し、入力音高を推定 (推定した音高はオクターブ 3,4)
- step 2. 推定した音高のうち、オクターブ 3 の音高をオクターブ 4 の音高に変換 (例: D3 → D4, A4)
- step 3. 変換したオクターブ 4 の音高に対応するくし形フィルタを縦続に接続
- step 4. 縦続接続したフィルタから入力音高を推定 (推定した音高はオクターブ 4)

7 段縦続接続くし形フィルタ推定される音高はオクターブ 3,4 であるため、オクターブ 4 の音高に変換する必要がある。

### 3.3 後段のくし形フィルタに対応した音高の推定

縦続接続くし形フィルタの音高の推定では、くし形フィルタの振幅特性により、低域の雑音成分が増加し、12 段を 7 段にしても、依然として後段のくし形フィルタに対応する音高の推定が困難となる問題が生じる。そこで、後段のくし形フィルタに対応する音高を推定するために、step1 の 7 段縦続接続くし形フィルタでの音高推定を、次の 2 つの手法を用いて補正を行う。

1 つめの手法として、HPF に信号を通すことで、低域の

雑音成分による出力値への影響を減少できる。ここで、推定対象がオクターブ 4 の音であるため、C4: 261.62Hz よりも周波数が低い帯域の信号を除去し、音高を推定する。

2 つめの手法として、最後段のくし形フィルタに対応する音高の推定を行うために、7 段縦続接続くし形フィルタでの推定を 2 回行う。そのため、3.2 の step1 が以下に示す手順となる。

- 1) 7 段縦続接続くし形フィルタで和音を推定 (仮推定)
- 2) 前段へ移動していきくし形フィルタの並びを反転させる
- 3) もう一度推定を行う (本推定)

縦続接続数が最大となる。最後段のくし形フィルタに対応する音高が推定されない場合が多い。そこで、まず仮推定として 1 度音高を推定し、移動しなかつくし形フィルタの並びを反転させる。入力音高である可能性が高いフィルタが前段へと移動し、今まで最後段にあったフィルタが中盤へ移動となるため、今まで推定できていなかった最後段のくし形フィルタに対応する音高が推定できるようになる。

## 4. ピアノ 4 和音の音高推定実験

### 4.1 実験条件

提案法と後段のくし形フィルタに対応する音高の推定のための補正が和音の推定に有効であるか、実際にピアノ和音の推定を行う。本実験では、RWC の楽音データベースのオクターブ 4 のピアノ楽音<sup>9)</sup> から作成したピアノ 4 和音 (495 通り) を入力信号として使用する。音高推定には音の鳴り始めから 0.6s の範囲のサンプルを用い、0.1s の観測窓を 0.01s ずらしながら推定を行う。なお、くし形フィルタは、対応する音高の周波数が低い順に縦続接続している (C3 → F#4 の順)。また、各和音の推定率は

$$\text{推定率} = \frac{\text{正答数}}{\text{推定回数}} \times 100 [\%] \quad (7)$$

で求める。また、各 4 和音 (495 通り) の音高推定率の平均をとったものを 4 和音の平均音高推定率とする。

### 4.2 7 段縦続接続くし形フィルタによる推定結果

7 段縦続接続くし形フィルタを用いた音高推定結果の平均音高推定率を表 2 に示す。この結果は、3.2 節の step1 の結果であり、表の 1/4~4/4 は 4 和音中の何音推定できているかを表し、パワー比とは 7 段くし形フィルタのみ、その他はこれに HPF、仮推定などの処理を追加した場合の音高推定結果を表している。また、7 段縦続接続くし形フィルタによる音高推定結果の一例を図 10,11 に示す。図 10 は 7 段縦続接続くし形フィルタによる音高推定結果であり、図 11 は HPF と仮推定の処理を追加した場合の音高推定結果である。図の横軸は時間、縦軸は推

定した音高を表しており、各時間毎の推定結果をプロットしている。また、背景が塗りつぶしてある部分が正しい音高である。入力音は次の通りである。

- 0 ~ 1s F# + A + A# + B
- 1 ~ 2s E + F + F# + A#
- 2 ~ 3s C# + D + E + G
- 3 ~ 4s D + F# + G + A
- 4 ~ 5s E + F# + G + G#
- 5 ~ 6s F# + G + G# + A

表 2 より、提案法を用いることで、7 段縦続接続くし形フィルタで 60% の推定率を得ることができた。また、HPF・仮推定の手法を用いることで、推定率を上げることができ、両方を用いることで推定率が約 30% 上がることを確認できる。このことから、今回用いた HPF と仮推定の手法は後段のくし形フィルタに対応する音高の推定に有効であるといえる。

図 10, 11 の推定例を見ると、HPF と仮推定を用いることで、推定のばらつきが全体的に少なくなっていることがわかる。また、フィルタの最後段である F# の推定は、図 10 では F と誤って推定しているが、図 11 だと最後段の F# が推定できている。7 段縦続接続くし形フィルタの推定で、最も推定が誤っていたのは最後段の推定であり、今回の補正でその最後段の音高を推定しているため、推定率が上昇したと考える。

しかし、完全に推定できているわけではなく、3s ~ 4s の間では、G が 1 音推定できていない。この時の推定の過程を図 12 に示す。丸で囲んだ音に入力音が含まれている。仮推定の時点では、最後段である F# の音高の代わりに F を推定し、次に D を推定している。F# を推定出来ないのは予想内だが、この時に G の音高が推定できていない。これは、F のフィルタで F# の音高を削った時のパワー比の減少が、G の音高を消去した時のパワー比減少よりも大きかったためだと考える。そのため、移動していないフィルタの並びを反転した時に、推定できていなかった音高 G が最後段へ移動してしまい、本推定でまた最後段の音高を推定できず、推定を誤っている。

7 段縦続接続くし形フィルタでは、縦続接続数を減らすためにオクターブ 3 のフィルタを利用しているが、正確に倍音成分としてオクターブ 4 の音を 2 つ消せるわけではなく、零点が少しずれているため、完全に除去しきれていない音が存在する。そのため、対応するくし形フィルタ以外のフィルタで信号が削られた際に、よりパワー比が減少する、あるいは、対応するくし形フィルタで信号を消去した際に、別の成分が増幅されて、パワー比があまり減少しないという場合があるため、このような問題が生じてしまう。

#### 4.3 4 和音の推定結果

7 段縦続接続くし形フィルタを用いて推定した音高を

表 2 4 和音の平均音高推定率 (7 段) [%]

正答数	パワー比	HPF	仮推定	HPF + 仮推定
1/4	99.87	99.98	99.91	99.99
2/4	98.60	99.47	98.88	99.74
3/4	92.08	96.44	94.72	98.48
4/4	60.09	69.46	72.71	87.38

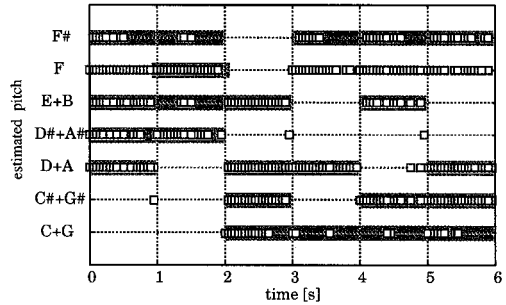


図 10 7 段縦続くし形フィルタのみの音高推定結果

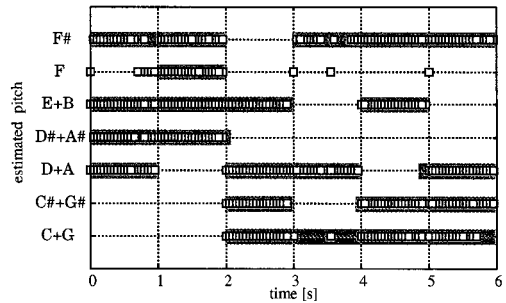


図 11 7 段縦続くし形フィルタ + HPF・仮推定の音高推定結果

表 3 4 和音の平均音高推定率 (12 音) [%]

正答数	パワー比	HPF + 仮推定
1/4	96.42	98.80
2/4	87.51	97.17
3/4	69.96	92.94
4/4	40.08	76.71

オクターブ 4 に対応させ、もう一度縦続接続くし形フィルタで音高を推定する。この処理は 3.2 節の step4 であり、最終的なピアノ 4 和音の推定率を求める。ピアノ 4 和音の平均音高推定率を表 3 に示し、音高推定結果の一例を図 13 に示す。入力音は前節と同じである。

表 3 より、7 段縦続接続くし形フィルタのみでの推定結果 40% が、HPF・仮推定の補正を行うことで、ピアノ 4 和音の平均音高推定率は 77% まで向上する。以前検討していた、並列構成くし形フィルタを用いた手法においては、4 和音の推定率は約 60% であったため、従来の手法よりも提案した手法を用いることで、推定率を高くすることができた。

ここで、図 13 の推定結果例は、図 11 の結果を基にし

input : D F# G A

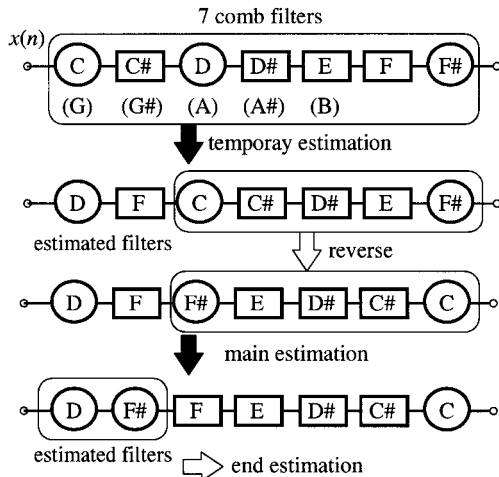


図 12 3s~4s 間の誤推定

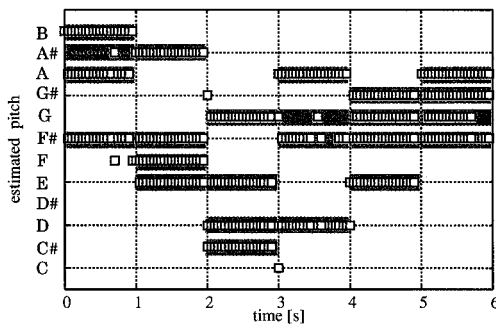


図 13 4 和音の音高推定結果 (HPF+仮推定)

ているので、3s~4sの間ではGが推定できない。step1での音高の推定を誤っていると、それ以上の推定結果は得られないため、7段縦続接続くし形フィルタでの推定精度をまだ上げる必要がある。逆に、0s~1s間の推定では、A#が1音推定できていない。3.2節のstep3で、オクターブ4の音高に対応するくし形フィルタを縦続接続する際、7段縦続接続くし形フィルタで音高を正確に推定していても、縦続数は4~8となる。そのため、接続数の増加により、後段の推定を誤る場合がある。しかし、この例の場合は、BとAの音高に対応するくし形フィルタで十分にA#の信号成分を消してしまったために、A#の信号を推定できなかったと考える。したがって、パワー比以外の零出力検出法についても検討を行う必要がある。

## 5. まとめ

本稿では、7段縦続接続くし形フィルタを用いた多和音(4和音)の音高推定法について検討を行った。実際にピアノ4和音を推定し、従来の並列方式の60%の推定率を、HPFと仮推定の補正を加えることで約77%まで上げることができた。しかし、完全に入力音の音高を推定できているわけではないので、まだ推定アルゴリズムについて検討する必要がある。

今後は、推定の精度を上げるとともに、推定対象となる和音を増やし、音域を広げて提案法の性能を明らかにする。

## 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) A. Sterian and G.H. Wakefield, "Music transcription systems: From sound to symbol," Proc. AAAI-2000 Workshop on AI and Music,2000.
- 3) 井口征士, "音楽情報の処理-電子計算機を用いた自動採譜," 計測自動制御学会論文集, vol.19,no.3,pp.314-319,1980.
- 4) 高澤嘉光, "計算機による採譜," 情報処理, vol.29,no.6,pp.593-598,1988.
- 5) 三輪多恵子, 田所嘉昭, 斎藤 努, "零出力に注目したくし形フィルタによる音階検出," 電学論 (C), vol.J118-C,no.1,pp.57-64,1998.
- 6) 山口 満, 三輪多恵子, 田所嘉昭, "並列構成くし形フィルタと特異値分解による多重唱の音高推定," 信学論 (D-II), vol.J87-D-II,no.4,pp.1020-1029, April, 2004.
- 7) 森田健夫, 山口 満, 田所嘉昭, "並列構成くし形フィルタの出力値に注目した採譜のための音高推定法," 信学論 (D-II), vol.J87-D-II,no.12,pp.2271-2279, Dec, 2004.
- 8) 松山大仁郎, 夏井雅典, 田所嘉昭, "7段縦続接続くし形フィルタによる多和音の音高推定法の検討," 電気関係学会東海支部連合大会,2006.
- 9) 後藤真孝, 橋口博樹, 西村拓一, "RWC 研究用音楽データベース: 音楽ジャンルデータベースと楽器音データベース," 情処研報, 2002-MUS-45, pp.19-26, 2002.