

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5011526号  
(P5011526)

(45) 発行日 平成24年8月29日 (2012. 8. 29)

(24) 登録日 平成24年6月15日 (2012. 6. 15)

(51) Int. Cl.	F I	
<b>G 1 O G</b> 3/04 (2006. 01)	G 1 O G	3/04
<b>G 1 O H</b> 1/00 (2006. 01)	G 1 O H	1/00 B
<b>G 1 O H</b> 1/38 (2006. 01)	G 1 O H	1/38 Z
<b>G O 1 H</b> 3/00 (2006. 01)	G O 1 H	3/00 B
<b>G 1 O L</b> 11/04 (2006. 01)	G 1 O L	11/04

請求項の数 2 (全 8 頁)

(21) 出願番号	特願2006-227773 (P2006-227773)	(73) 特許権者	304027349
(22) 出願日	平成18年8月24日 (2006. 8. 24)		国立大学法人豊橋技術科学大学
(65) 公開番号	特開2008-52023 (P2008-52023A)		愛知県豊橋市天伯町雲雀ヶ丘 1-1
(43) 公開日	平成20年3月6日 (2008. 3. 6)	(72) 発明者	田所 嘉昭
審査請求日	平成21年8月24日 (2009. 8. 24)		愛知県豊橋市天伯町雲雀ヶ丘 1-1
			国立大学法人豊橋技術科学大学内
		(72) 発明者	夏井 雅典
			愛知県豊橋市天伯町雲雀ヶ丘 1-1
			国立大学法人豊橋技術科学大学内
		(72) 発明者	松山 大仁郎
			愛知県豊橋市天伯町雲雀ヶ丘 1-1
			国立大学法人豊橋技術科学大学内

最終頁に続く

(54) 【発明の名称】 多和音の音名と音高推定手法

(57) 【特許請求の範囲】

【請求項 1】

音名推定のための処理方法であって、

対象音域の最低オクターブにある 1 2 音の各音を消去するくし形フィルタを音名の高い順に並べて縦続接続して初期状態とする第一の工程と、

縦続接続されたくし形フィルタに対象多和音を入力し、前記各くし形フィルタの入出力周波数比の変化を検出する第二の工程と、

前記第二の工程により検出された入出力周波数比の変化が大きく、縦続接続の末尾に近いくし形フィルタを特定する第三の工程と、

前記第三の工程により特定されたくし形フィルタを、縦続接続の先頭に移動させて縦続順序を変更する第四の工程と、

を含み、第三の工程により特定されたくし形フィルタが最初に特定されたくし形フィルタと一致するまで、前記第二の工程から第四の工程を繰り返すことを特徴とする音名推定のための処理方法。

【請求項 2】

請求項 1 に記載の処理方法において、

前記第三の工程により特定されたくし形フィルタについて、該くし形フィルタの次数（遅延数）を半減させる第五の工程と、

前記第五の工程により次数が半減されたくし形フィルタの入出力周波数比の変化または入出力振幅比を検出する第六の工程と、

前記第六の工程により、入出力周波数比の変化が生じないか、または、入出力振幅比が1以上になるまで、前記第五および第六の工程を繰り返すことを特徴とする音高推定のための処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、楽音を解析し楽譜として出力する採譜システムの実現において、出力波形における振幅の変化や周波数の変化に注目して多和音を対象とした音高推定手法に関するものである。

【背景技術】

【0002】

音楽信号を楽譜に書き起こす処理のことを採譜と呼ぶ。従来、採譜は、音楽に関する知識や経験を持つ専門家にしか行えないものであった。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を実現するためには、楽音の音の高さ（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、音高推定とは主に基本周波数（ピッチ）を推定する処理のことを指す。

【0003】

従来技術として、マイクなどの入力装置、あるいはCD再生などによって入力された音響信号を楽譜に直す採譜システム、例えば、ヤマハ社のXGワークスV2.0、USAワ  
イルドキャットキャニオン社のオートスコアなどがある。これらのほとんどは独唱や単  
独の楽器によって演奏される単音を対象にしたものであり、周波数特性の異なる複数の音源  
によって構成された和音までを扱うことが可能なものは、いまだ実用化されていない。単  
音に関する特許文献として、特開平05-108062号公報、特開平05-12766  
8号公報、特開平05-165464号公報、特開平09-072779号公報がある。

【0004】

一方、これまでの和音を対象とした音高推定のための技術として、和音の採譜に関する  
特許文献では、特開平04-278544号公報、特開平05-1006602号公報、  
特開平05-173557号公報、特開2005-202354A号公報がある。しかし  
、これらに開示されている技術では、10和音などの多和音に対応することは難しい。

【0005】

また、非特許文献1では、楽音のすべての成分を消去することができるくし形フィルタ  
を使用し、1オクターブにある12音の各音に対応する12個のくし形フィルタを用いる  
ことで、何和音でもその各音の音高を推定する方法が提案されている。しかし、使用され  
ている楽音はMIDI音源（電子楽器）に対する多和音の音高推定手法に関するものであ  
る。この手法を実楽器音源データ（RWC音楽データベースに収録、非特許文献3を参照  
）に使用した場合、実楽器のもつ非定常性のため、フィルタの振幅に注目した音高推定は  
困難になる。

【0006】

最後に、非特許文献2では、非特許文献1で明らかになった実楽器の非定常性に対応す  
るための音高推定法が提案されている。並列構成くし形フィルタの入出力振幅比を基本に  
して音高を推定する方法である。しかしながら、和音の数だけのくし形フィルタが接続さ  
れることになり、和音数の増加とともに音高を推定する能力は低下する。

【非特許文献1】三輪、田所、斉藤、「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. 1965-1  
974、1998

【非特許文献2】森田、山口、田所、「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、Vol. J87D-II、No. 12、pp. 2271  
- 2279、2004

【非特許文献3】後藤、橋口、西村、「RWC研究用音楽データベース：音楽ジャンルデ

10

20

30

40

50

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

【発明の開示】

【発明が解決しようとする課題】

【0007】

楽音は、基本周波数とその倍音成分から構成されるため、和音の場合、各音の成分が非常に複雑な関係になり、信号処理の一般的手法である離散フーリエ変換を基本にしたアルゴリズムでは、一般に高精度の推定は3和音程度までの音高推定である。しかし、実際の演奏では4和音以上も存在することが多い。例えば、ピアノ演奏では、10本の指を使えば10和音も生成できる。そのため4和音以上の多和音についても、その各音高を精度良く推定する手法を開発することは、実用的な採譜システムを実現するために是非必要とされる技術である。

10

【0008】

そこで、本発明で解決する課題は、縦続接続くし形フィルタシステムを用い、フィルタ出力における振幅の変化や周波数の変化から多和音の音高を推定する手法を提供することである。

【課題を解決するための手段】

【0009】

実楽器音の多和音を縦続接続くし形フィルタで処理する時、くし形フィルタの入出力振幅比だけでなく、入出力周波数比にも注目する点が本発明のポイントである。楽音の非定常性が縦続接続くし形フィルタの周波数特性のために増幅され、振幅の大小だけでは音高推定が不可能である。しかし、楽音の基本となる各周波数成分がくし形フィルタでほぼ除かれた後に、非定常成分の低周波成分が出力される。このため、くし形フィルタの入出力周波数比に注目し、その周波数比の大きな変化点を示すくし形フィルタから、対応する音名(C、C#、・・・、Bなど)が推定され、さらに、そのくし形フィルタの次数を各オクターブに対応するように変化させ、その出力波形の変化(振幅変化か周波数変化)からその音のオクターブが推定され、多和音のうちの1音の音高が推定できる。

20

【0010】

このようにして推定した1音に対応したくし形フィルタを縦続接続くし形フィルタの先頭に移動させ、再度、波形の変化、すなわち振幅比や周波数比に注目して残りの音名(オクターブ推定の処理を含めると音高)の一つを推定し、再度このくし形フィルタを縦続接続くし形フィルタの先頭に移動させる。この操作を繰り返して、最終的に最初に移動させたくし形フィルタでの波形の変化に大きな変化が生じた時点で音高推定操作が終了する。そして、移動させたくし形フィルタから多和音の各音高が推定される。

30

【0011】

対象音域の最低オクターブ中に存在する12音に対応するくし形フィルタを縦続接続し、その各くし形フィルタの入出力振幅比、周波数比の変化を測定し、その変化の大きいくし形フィルタを順次、縦続接続くし形フィルタの先頭に移動させる。なお、縦続接続するくし形フィルタの数は、対象音域の最低オクターブよりひとつ低いオクターブに対応するくし形フィルタを使用すれば、等価的に7個で済まされることも可能である。

40

【発明の効果】

【0012】

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簡単化を図ることができる。

【発明を実施するための最良の形態】

【0013】

1オクターブ12音に対応するくし形フィルタを図1のように縦続接続したシステムに

50

ついて考える。このシステムは、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。ここで、このフィルタの最終出力の振幅特性は図 2 のようになる。原理的には、入力和音中の全ての音が消去された段階でくし形フィルタの出力は零となるが、実際の楽音では非定常性、周波数変動のため、完全な零出力とならず、むしろ図 2 の特性により、低域の周波数成分が増大し、振幅の大小での音高推定は困難になる。

【 0 0 1 4 】

そこで、振幅値に注目するのではなく、時間波形（周波数）の変化に注目した方法を提案する。縦続接続くし形フィルタに信号を入力し、和音の周波数成分が全て除去されると、くし形フィルタの出力は零出力にはならないが、その時点で時間波形は楽音の非定常性により、周波数変動分の低い周波数の出力波形になる。この時間波形の変化に注目し、和音の音高推定を行う。

10

【 0 0 1 5 】

例として、図 3 に 5 和音（D + F + G + A + B）を入力した時の縦続接続くし形フィルタの各フィルタの出力を示す。各フィルタの出力波形を見ると、B から D # までのフィルタの出力波形に対し、音高 D のフィルタまで通した時点で時間波形（周波数）は、低い周波数の出力波形になっているのがわかる。この変化により、入力和音に D が含まれていると推定でき、B から D までの間の音高の組み合わせで入力和音が構成されているとも考えられる。

【 0 0 1 6 】

以上のことより、提案する音高推定アルゴリズムは図 4 のようになる。

20

【 0 0 1 7 】

（くし形フィルタによる音高推定の基本原理）

音高  $p$  の基本周波数を  $f_p$  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  $H_p(z)$  は、数 1 で表され、その出力  $y_p(n)$  は、数 2 に示す減算で実現される。

【 0 0 1 8 】

【 数 1 】

$$H_p(z) = 1 - z^{-N_p}$$

30

【 0 0 1 9 】

【 数 2 】

$$y_p(n) = x(n) - x(n - N_p)$$

【 0 0 2 0 】

ここで、 $N_p$  は遅延数であり、サンプリング周波数を  $f_s$  とすると  $N_p$  は、数 3 で求められる。

【 0 0 2 1 】

【 数 3 】

$$N_p = \lceil f_s / f_p \rceil \quad [\text{integer of rounding}]$$

40

【 0 0 2 2 】

図 5 にオクターブ 4 の音名 E の音楽信号（基本周波数は 329.63 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。

【 実施例 】

【 0 0 2 3 】

対象音域をオクターブ 3 以上としたときは、オクターブ 3 の各音名に対応するくし形フィルタを図 1 のように 12 個縦続接続し、このシステムに多和音を入力させる。この縦続

50

接続くし形フィルタの振幅特性は、図2のようになり、オクターブ3以下の音域の振幅ゲインが大きくなっている。そのため、多和音の各周波数成分が除かれた後には、理想的にはくし形フィルタの出力振幅は零になるはずであるが、実楽器の多和音の非定常性成分（低周波成分）が出力されることになる。その出力振幅は後段のくし形フィルタになるほど大きくなり、入出力振幅比での音名推定は困難になる。しかし、くし形フィルタの入出力周波数比に注目することで、多和音がすべて除かれたかどうかを判断できる。

#### 【0024】

入出力周波数比に注目して多和音の各音名を推定する例を図6に示す。図6では、左側が縦続接続くし形フィルタの移動を示し、対応する右側は和音の全周波数成分が消去される前後の時間波形をinputとoutputとしてそれぞれ示している。多和音の例として、5和音（オクターブ4：D + F + G + A + B）を考える。縦続接続くし形フィルタ（オクターブ4の12音に対応するくし形フィルタから構成）は、音名の高い順に並べている。より確実な推定をするためには、並べ方を逆にした二つの縦続接続くし形フィルタを使用するとよい。

#### 【0025】

最初に入出力周波数比が大きく変わるのは、音名Dを除くくし形フィルタである。なお、入出力振幅比(出力/入力)も1以下になっている。このDに対応するくし形フィルタを縦続接続くし形フィルタの先頭に移動させ、2回目の入出力周波数比が大きく変わるくし形フィルタを見つける。今度は音名Fを除くくし形フィルタで検出される。このFのくし形フィルタを先頭に移動させる。3回目の入出力周波数比の大きな変化は、音名Gに対応するくし形フィルタである。このGのくし形フィルタを先頭に移動させる。4回目の入出力周波数比の大きな変化点は、音名Aのくし形フィルタである。Aのくし形フィルタを先頭に移す。そして、5回目の入出力周波数比の大きな変化するくし形フィルタは音名Bのくし形フィルタである。このBに対応するくし形フィルタを先頭に移動し、6回目の入出力周波数比の大きな変化するくし形フィルタを検出する。今度は音名Dに相当するくし形フィルタであり、これは最初に先頭に移動されたくし形フィルタであるので、ここで音名推定の操作は終了する。

#### 【0026】

このようにして、各操作で先頭に移動されたくし形フィルタに相当する音名が、入力多和音の各音名に対応する。すなわち、入力されたD、F、G、A、Bが推定されている。

#### 【0027】

次に、これらの各音名のオクターブの推定は、各音名が推定された時点で、そのくし形フィルタの次数を2分の1、4分の1、・・・と変化させて、入出力周波数比の大きな変化が生じなく、または入出力振幅比(出力/入力)が1以上になる、次数の手前の次数が対応するオクターブとして決定される。そして、音名とオクターブから音高が決定される。なお、高調波成分のため、入出力周波数比（ゼロクロス数、ピーク数）の判定が入出力波形そのものから難しい場合は、平均化フィルタや自己相関関数の処理で明確にできる。この例の場合は、オクターブ推定でくし形フィルタの次数を2分の1にしたときに入出力周波数比の大きな変化が生じなくなる。よって、これらのオクターブは4と決定される。

#### 【産業上の利用可能性】

#### 【0028】

本発明により、和音の数に制限されずに多和音の各音高を推定することができ、本格的な採譜システムを実現できる。これにより、音楽配信ビジネス、カラオケ産業での楽曲検索、民族音楽の記録の効率化、即興音楽の記録化、作曲作業の能率化および音楽教育への利用など、その応用範囲は多岐にわたる。音楽配信ビジネス（通信カラオケ、携帯電話やPHSの着信メロディ）では、新曲をいち早く楽譜化することが必要であり、これを人手ではなく計算機による自動採譜システムで行なうことが熱望されている。この自動採譜システムとして4和音以上の多和音の精度良い音高推定が可能な本格的な採譜システムが実現されれば、この業界に多大な貢献をすることが可能である。

#### 【0029】

また、楽音のリアルタイム認識によるセッションシステム（相手の演奏の理解に基づいた演奏応答システム）のようなインタラクティブ音楽コンテンツの実現、さらには和音の組み合わせの多さを利用した個人認証システム（例えば、特許文献1を参照）への応用など考えられる。携帯電話の音源を利用した個人認証システムでは、採譜システムによる音高推定能力により、その性能（秘話性）が決定され、本発明による本格的な採譜システムは、そのための有望な技術になり、これに関係する産業界、IT社会での本発明の使用が期待される。

【特許文献1】特開2001-318894号公報

【図面の簡単な説明】

【0030】

【図1】くし形フィルタ（オクターブ3）の12段縦続接続によるシステム構成

【図2】 $H_p(z)$ （ $p$ は1から12）を12段接続した縦続接続くし形フィルタの振幅特性

【図3】5和音（D + F + G + A + B）を入力したときの縦続接続くし形フィルタの各出力

【図4】図1のように構成されるくし形フィルタシステムの処理アルゴリズム

【図5】オクターブ4の音名Eの周波数特性とくし形フィルタの振幅特性

【図6】本発明の実施例に係る多和音（オクターブ4：D + F + G + A + B）の音名推定過程

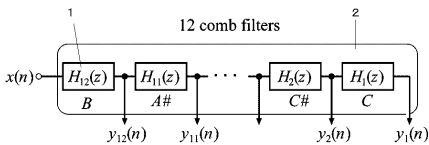
【符号の説明】

【0031】

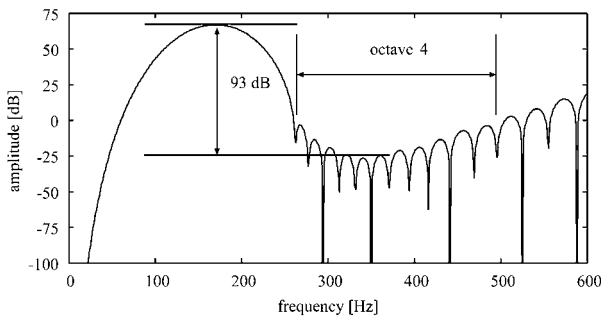
1  $H_p(z) = 1 - z^{-N_p}$  の特性を持つくし形フィルタ

2 くし形フィルタを12段縦続接続構成にしたくし形フィルタシステム

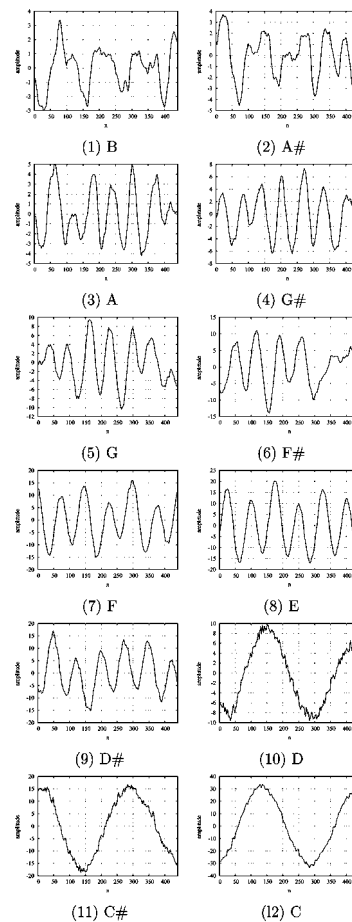
【図1】



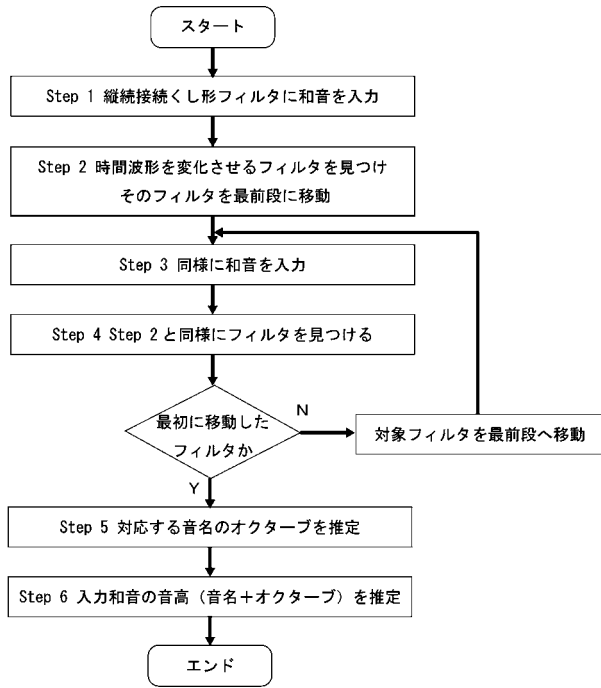
【図2】



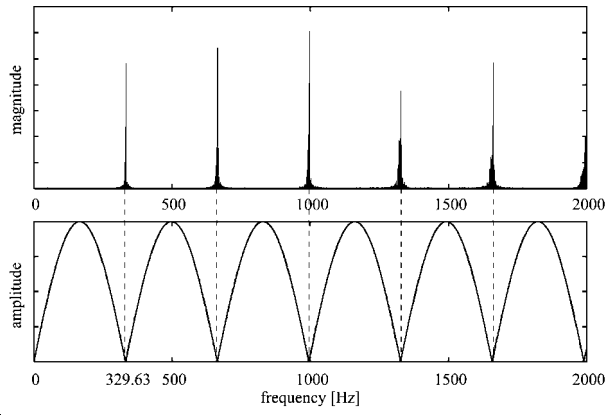
【図3】



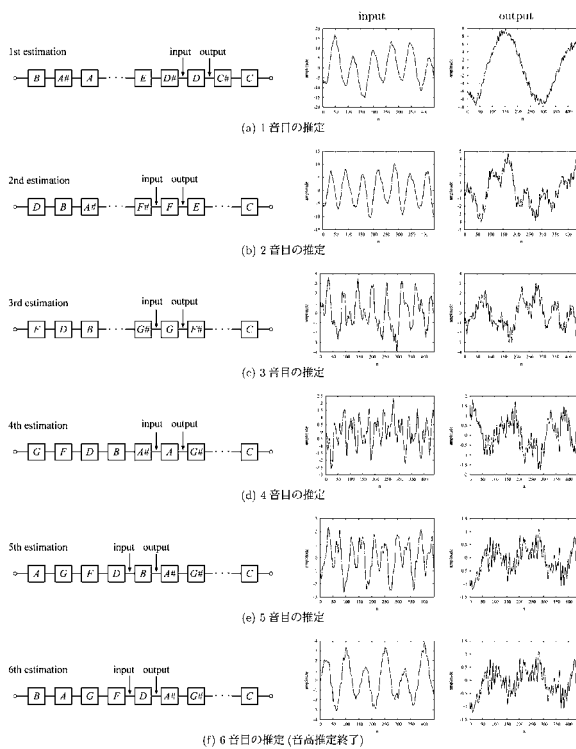
【 図 4 】



【 図 5 】



【 図 6 】



---

フロントページの続き

審査官 毛利 太郎

- (56)参考文献 特開平06 - 149242 (JP, A)  
特開平10 - 011066 (JP, A)  
特開平05 - 173557 (JP, A)  
特開平09 - 083299 (JP, A)  
特開2004 - 349739 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G10G 3/00 - 3/04  
G10H 1/00 - 7/12  
G10L 11/00 - 11/06  
G01H 3/00