

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

W02010/095622

発行日 平成24年8月23日 (2012.8.23)

(43) 国際公開日 平成22年8月26日 (2010.8.26)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G 1 0 L 21/02 (2006.01)</b>	G 1 0 L 21/02 3 0 1 B	5 D 3 7 8
<b>G 1 0 H 1/00 (2006.01)</b>	G 1 0 L 21/02 2 0 1 D	
	G 1 0 H 1/00 B	

審査請求 未請求 予備審査請求 未請求 (全 47 頁)

出願番号 特願2011-500614 (P2011-500614)	(71) 出願人 504132272 国立大学法人京都大学 京都府京都市左京区吉田本町36番地1
(21) 国際出願番号 PCT/JP2010/052293	(74) 代理人 100091443 弁理士 西浦 ▲嗣▼晴
(22) 国際出願日 平成22年2月16日 (2010.2.16)	(72) 発明者 安部 武宏 京都府京都市左京区吉田本町 国立大学法人京都大学大学院情報学研究科内
(31) 優先権主張番号 特願2009-34664 (P2009-34664)	(72) 発明者 安良岡 直希 京都府京都市左京区吉田本町 国立大学法人京都大学大学院情報学研究科内
(32) 優先日 平成21年2月17日 (2009.2.17)	(72) 発明者 糸山 克寿 京都府京都市左京区吉田本町 国立大学法人京都大学大学院情報学研究科内
(33) 優先権主張国 日本国 (JP)	

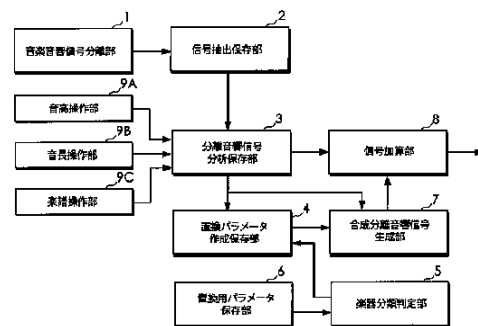
最終頁に続く

(54) 【発明の名称】 音楽音響信号生成システム

(57) 【要約】

既存の音楽音響信号中の音色を任意の音色に変えることができる音楽音響信号の音色変更システムを提供する。分離音響信号分析保存部3に保存された、第1の種類の楽器の単音ごとのn次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、置換用パラメータデータ保存部6に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音のn次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成する。合成分離音響信号生成部7は、倍音ピーク・パラメータを除く他のパラメータと置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。

【図1】



- 1 Music acoustic signal separating unit
- 2 Signal extraction/storage unit
- 3 Separated acoustic signal analysis/storage unit
- 4 Replacement parameter generation/storage unit
- 5 Musical Instrument Identification unit
- 6 Replacement parameter storage unit
- 7 Synthesized separated acoustic signal generating unit
- 8 Signal adding unit
- 9A Sound pitch manipulating unit
- 9B Sound length manipulating unit
- 9C Musical note manipulating unit

## 【特許請求の範囲】

## 【請求項 1】

第 1 の種類の楽器から発生した楽器音の音響信号を含む音楽音響信号から抽出した、前記第 1 の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとに保存し且つ残差音響信号を保存する信号抽出保存部と、

前記単音ごとの分離音響信号を、少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータと  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析して保存する分離音響信号分析保存部と、

前記第 1 の種類の楽器とは異なる第 2 の種類の楽器から発生した楽器音の音響信号から作成した、前記分離音響信号に含まれる全ての単音に対応する前記第 2 の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要な、前記第 2 の種類の楽器から発生した前記複数の単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータを保存する置換用パラメータ保存部と、

前記分離音響信号分析保存部に保存された、前記第 1 の種類の楽器の単音ごとの前記  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記置換用パラメータ保存部に保存された、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより作成した置換倍音ピーク・パラメータを保存する置換パラメータ作成保存部と、

前記分離音響信号分析保存部に保存された前記倍音ピーク・パラメータを除く他のパラメータと前記置換用パラメータ保存部に保存された前記置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成する合成分離音響信号生成部と、

前記合成分離音響信号と前記残差音響信号とを加算して、第 2 の種類の楽器から発生した楽器音を含む音楽音響信号を出力する信号加算部とからなる音楽音響信号生成システム。

## 【請求項 2】

第 1 の種類の楽器から発生した楽器音を含む音楽音響信号から抽出した、前記第 1 の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとに保存し且つ残差音響信号を保存する信号抽出保存部と、

単音ごとの前記分離音響信号を、少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータと  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析して保存する分離音響信号分析保存部と、

前記第 1 の種類の楽器とは異なる第 2 の種類の楽器から発生した楽器音の音響信号から作成した、前記分離音響信号に含まれる全ての単音に対応する前記第 2 の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要な、前記第 2 の種類の楽器の前記複数の単音ごとの  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータ及び  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを保存する置換用パラメータ保存部と、

前記分離音響信号分析保存部に保存された、前記第 1 の種類の楽器の単音ごとの前記  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記置換用パラメータ保存部に保存された、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより作成した置換倍音ピーク・パラメータを保存し、且つ前記分離音響信号分析保存部に保存された、前記第 1 の種類の楽器の単音ごとの前記  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、前記置換用パラメータ保存部に保存された、前記第 1 の種類の楽器の単音に対

10

20

30

40

50

応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータと置き換えることにより作成した置換パワーエンベロープ・パラメータを保存する置換パラメータ作成保存部と、

前記分離音響信号分析保存部に保存された前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを除く他のパラメータと前記置換パラメータ作成保存部に保存された前記置換倍音ピーク・パラメータ及び前記置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成する合成分離音響信号生成部と、

前記合成分離音響信号と前記残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力する信号加算部とからなる音楽音響信号生成システム

10

### 【請求項3】

第1の種類の楽器から発生した楽器音を含む音楽音響信号から抽出した、前記第1の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとに保存し且つ残差音響信号を保存する信号抽出保存部と、

前記単音ごとの分離音響信号を、少なくとも $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータと $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析して保存する分離音響信号分析保存部と、

前記第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から作成した、前記音楽音響信号に含まれる全ての単音に対応する前記第2の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要な、前記第2の種類の楽器から発生した前記複数の単音ごとの $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータ及び次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを保存する置換用パラメータ保存部と、

前記第1の種類の楽器と前記第2の種類の楽器とが、同じ楽器分類に属するか否かを判定する楽器分類判定部と、

前記分離音響信号分析保存部に保存された、前記第1の種類の楽器の単音ごとの前記 $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記置換用パラメータ保存部に保存された、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより作成した置換倍音ピーク・パラメータを保存し、且つ前記分離音響信号分析保存部に保存された、前記第1の種類の楽器の単音ごとの前記 $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、前記置換用パラメータ保存部に保存された、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータと置き換えることにより作成した置換パワーエンベロープ・パラメータを保存する置換パラメータ作成保存部と、

前記楽器分類判定部が、前記第1の種類の楽器と前記第2の種類の楽器とが、同じ楽器分類に属すると判定したときには、前記分離音響信号分析保存部に保存された前記倍音ピーク・パラメータを除く他のパラメータと前記置換パラメータ作成保存部に保存された前記置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成し、前記楽器分類判定部が、前記第1の種類の楽器と前記第2の種類の楽器とが、異なる楽器分類に属すると判定したときには、前記分離音響信号分析保存部に保存された前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを除く他のパラメータと前記置換パラメータ作成保存部に保存された前記置換倍音ピーク・パラメータ及び前記置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成する合成分離音響信号生成部と、

前記合成分離音響信号と前記残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力する信号加算部とからなる音楽音響信号生成システム

20

30

40

50

。

## 【請求項 4】

前記分離音響信号分析保存部は、第 1 の種類の楽器の単音ごとの非調波成分分布パラメータを保存する機能を更に備えており、

前記置換用パラメータ保存部は、前記第 2 の種類の楽器から発生した楽器音の音響信号の前記複数種類の単音ごとの非調波成分分布パラメータを保存する機能を更に備えており、

前記置換パラメータ作成保存部は、前記分離音響信号分析保存部に保存された、前記第 1 の種類の楽器の単音ごとの前記非調波成分分布パラメータを、前記置換用パラメータ保存部に保存された、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の前記非調波成分分布パラメータと置き換えることにより作成した置換非調波成分分布パラメータを更に保存し、

前記合成分離音響信号生成部は、前記分離音響信号分析保存部に保存された前記倍音ピーク・パラメータ、前記パワーエンベロープ・パラメータ及び前記非調波成分分布パラメータを除く他のパラメータと前記置換パラメータ作成保存部に保存された前記置換倍音ピーク・パラメータ、前記置換パワーエンベロープ・パラメータ及び前記非調波成分分布パラメータとを用いて、単音ごとの合成分離音響信号を生成する請求項 2 または 3 に記載の音楽音響信号生成システム。

## 【請求項 5】

前記置換用パラメータ保存部は、前記第 2 の種類の楽器から発生した楽器音の音響信号から得た複数種類の単音の分離音響信号を前記調波モデルにより表現する場合に必要なとなる、前記複数種類の単音ごとの少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータを分析して保存し、併せて前記複数種類の単音ごとの  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを保存するパラメータ分析保存部と、

前記パラメータ分析保存部に保存した前記複数種類の単音についての前記倍音ピーク・パラメータと前記パワーエンベロープ・パラメータとに基づいて、前記音楽音響信号に含まれる全ての単音に対応する前記第 2 の種類の楽器から発生する複数の単音のうち前記複数種類の単音以外の単音についての音響信号を前記調波モデルにより表現する場合に必要なとなる前記第 2 の種類の楽器の前記複数の単音ごとの前記倍音ピーク・パラメータを補間法を用いて生成して保存するパラメータ補間生成保存部とからなり、

前記パラメータ分析保存部は、分析により得られた前記  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、代表パワーエンベロープ・パラメータとして保存する請求項 2 または 3 に記載の音楽音響信号生成システム。

## 【請求項 6】

前記置換用パラメータ保存部は、前記複数種類の単音ごとの少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータ及び  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを分析して保存するパラメータ分析保存部と、

前記パラメータ分析保存部に保存した前記複数種類の単音についての前記倍音ピーク・パラメータと前記パワーエンベロープ・パラメータとに基づいて、前記音楽音響信号に含まれる全ての単音に対応する前記第 2 の種類の楽器から発生する複数の単音のうち前記複数種類の単音以外の単音についての音響信号を前記調波モデルにより表現する場合に必要なとなる前記第 2 の種類の楽器の前記複数の単音ごとの前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを補間法を用いて生成して保存するパラメータ補間生成保存部とからなる請求項 2 または 3 に記載の音楽音響信号生成システム。

## 【請求項 7】

前記置換用パラメータ保存部は、前記パラメータ分析保存部及び前記パラメータ補間生成保存部に保存されたデータに基づいて、前記第 2 の種類の前記複数の単音ごとの前記倍音ピーク・パラメータを音高依存特徴関数として保存する関数生成保存部をさらに備え、

前記置換パラメータ作成保存部は、前記第 2 の種類の楽器の前記単音の前記倍音ピーク

・パラメータに含まれる複数の倍音ピークを前記音高依存特徴関数から取得するように構成されている請求項 5 に記載の音楽音響信号生成システム。

【請求項 8】

前記音楽音響信号を含む混合音響信号から前記音楽音響信号を分離する音響信号分離部をさらに備えている請求項 1, 2 または 3 に記載の音楽音響信号生成システム。

【請求項 9】

前記音楽音響信号を含む混合音響信号から前記音楽音響信号を分離する音響信号分離部をさらに備えており、前記音楽音響信号以外の音響信号が前記残差音響信号中に含まれる請求項 1, 2 または 3 に記載の音楽音響信号生成システム。

【請求項 10】

前記音楽音響信号を含む混合音響信号から得た別の音楽音響信号から前記第 2 の種類の楽器の楽器音を取得する請求項 9 に記載の音楽音響信号生成変更システム。

【請求項 11】

前記調波モデルが、倍音構造の非調和性を組み込んだ調波モデルである請求項 1, 2 または 3 に記載の音楽音響信号生成システム。

【請求項 12】

前記分離音響信号分析保存部が分析する複数のパラメータには、音高に関する音高パラメータと音長に関する音長パラメータとが含まれており、

前記音高パラメータを操作する音高操作部と、前記音長パラメータを操作する音長パラメータ操作部をさらに備えている請求項 1, 2 または 3 に記載の音楽音響信号生成システム。

【請求項 13】

第 1 の種類の楽器から発生した楽器音を含む音楽音響信号から、前記第 1 の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとにそれぞれ抽出し且つ残差音響信号を抽出するステップと、

単音ごとの前記分離音響信号を、少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータと  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析するステップと、

前記第 1 の種類の楽器とは異なる第 2 の種類の楽器から発生した楽器音の音響信号から、前記音楽音響信号に含まれる全ての単音に対応する前記第 2 の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要となる、前記第 2 の種類の楽器の前記複数の単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータを作成するステップと、

前記第 1 の種類の楽器の単音ごとの前記  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成するステップと、

前記倍音ピーク・パラメータ除く他のパラメータと前記置換用パラメータ保存部に保存された前記置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成するステップと、

前記合成分離音響信号と前記残差音響信号とを加算して、第 2 の種類の楽器から発生した楽器音を含む音楽音響信号を出力するステップとをコンピュータが実施する音響信号生成方法。

【請求項 14】

第 1 の種類の楽器から発生した楽器音を含む音楽音響信号から、前記第 1 の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとにそれぞれ抽出し且つ残差音響信号を抽出するステップと、

単音ごとの前記分離音響信号を、少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク

10

20

30

40

50

・パラメータと $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ  
 ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析するステップと、

前記第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から、前記音楽音響信号に含まれる全ての単音に対応する前記第2の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要となる、前記第2の種類の楽器の前記複数の単音ごとの $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータ及び $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを作成するステップと、

前記第1の種類の楽器の単音ごとの前記 $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成し、且つ前記第1の種類の楽器の単音ごとの前記 $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域を、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域と置き換えることにより置換パワーエンベロープ・パラメータを作成するステップと、

前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを除く他のパラメータと前記置換倍音ピーク・パラメータ及び前記置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成するステップと、

前記合成分離音響信号と前記残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力するステップとをコンピュータが実施することを特徴とする音楽音響信号生成方法。

#### 【請求項15】

第1の種類の楽器から発生した楽器音を含む音楽音響信号から、前記第1の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとにそれぞれ抽出し且つ残差音響信号を抽出するステップと、

単音ごとの前記分離音響信号を、少なくとも $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータと $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析するステップと、

前記第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から、前記音楽音響信号に含まれる全ての単音に対応する前記第2の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要となる、前記第2の種類の楽器の前記複数の単音ごとの $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータ及び $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを作成するステップと、

前記第1の種類の楽器と前記第2の種類の楽器とが、同じ楽器分類に属するか否かを判定するステップと、

前記第1の種類の楽器の単音ごとの前記 $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記置換用パラメータ保存部に保存された、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成し、且つ前記第1の種類の楽器の単音ごとの前記 $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域を、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音の $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域と置き換えることにより置換パワーエンベロープ・パラメータを作成するステップと、

10

20

30

40

50

前記楽器分類判定部が、前記第1の種類の楽器と前記第2の種類の楽器とが、同じ楽器分類に属すると判定したときには、前記倍音ピーク・パラメータ除く他のパラメータと前記置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成し、前記楽器分類判定部が、前記第1の種類の楽器と前記第2の種類の楽器とが、異なる楽器分類に属すると判定したときには、前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを除く他のパラメータと前記置換倍音ピーク・パラメータ及び前記置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成するステップと、

前記合成分離音響信号と前記残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力するステップとをコンピュータが実施する音楽音響信号生成方法。

10

【請求項16】

第1の種類の楽器から発生した楽器音を含む音楽音響信号から、前記第1の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとにそれぞれ抽出し且つ残差音響信号を抽出するステップと、

単音ごとの前記分離音響信号を、少なくともn次倍音成分の相対強度を示す倍音ピーク・パラメータとn次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析するステップと、

前記第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から、前記音楽音響信号に含まれる全ての単音に対応する前記第2の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要となる、前記第2の種類の楽器の前記複数の単音のn次倍音成分の相対強度を示す倍音ピーク・パラメータを作成するステップと、

20

前記第1の種類の楽器の単音ごとの前記n次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記第1の種類の楽器の単音に対応する前記第2の種類の楽器の前記単音のn次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成するステップと、

前記倍音ピーク・パラメータ除く他のパラメータと前記置換用パラメータ保存部に保存された前記置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成するステップと、

30

前記合成分離音響信号と前記残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力するステップとをコンピュータを用いて実施するために前記コンピュータで用いられる音楽音響信号生成用コンピュータプログラム。

【請求項17】

第1の種類の楽器から発生した楽器音を含む音楽音響信号から、前記第1の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとにそれぞれ抽出し且つ残差音響信号を抽出するステップと、

単音ごとの前記分離音響信号を、少なくともn次倍音成分の相対強度を示す倍音ピーク・パラメータとn次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析するステップと、

40

前記第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から、前記音楽音響信号に含まれる全ての単音に対応する前記第2の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要となる、前記第2の種類の楽器の前記複数の単音ごとのn次倍音成分の相対強度を示す倍音ピーク・パラメータ及びn次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータ前記第1の種類の楽器から発生した楽器音の音響信号のみを含む作成するステップと、

50

前記第 1 の種類の楽器の単音ごとの前記 n 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の n 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成し、且つ前記第 1 の種類の楽器の単音ごとの前記 n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域を、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域と置き換えることにより置換パワーエンベロープ・パラメータを作成するステップと、

前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを除く他のパラメータと前記置換倍音ピーク・パラメータ及び前記置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成するステップと、

前記合成分離音響信号と前記残差音響信号とを加算して、第 2 の種類の楽器から発生した楽器音を含む音楽音響信号を出力するステップとをコンピュータを用いて実施するために前記コンピュータで用いられる音楽音響信号生成用コンピュータプログラム。

【請求項 18】

第 1 の種類の楽器から発生した楽器音を含む音楽音響信号から、前記第 1 の種類の楽器から発生した楽器音の音響信号のみを含む分離音響信号を単音ごとにそれぞれ抽出し且つ残差音響信号を抽出するステップと、

単音ごとの前記分離音響信号を、少なくとも n 次倍音成分の相対強度を示す倍音ピーク・パラメータと n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析するステップと、

前記第 1 の種類の楽器とは異なる第 2 の種類の楽器から発生した楽器音の音響信号から、前記音楽音響信号に含まれる全ての単音に対応する前記第 2 の種類の楽器から発生する複数の単音についての音響信号を前記調波モデルにより表現する場合に必要となる、前記第 2 の種類の楽器の前記複数の単音ごとの n 次倍音成分の相対強度を示す倍音ピーク・パラメータ及び n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを作成するステップと、

前記第 1 の種類の楽器と前記第 2 の種類の楽器とが、同じ楽器分類に属するか否かを判定するステップと、

前記第 1 の種類の楽器の単音ごとの前記 n 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、前記置換用パラメータ保存部に保存された、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の n 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成し、且つ前記第 1 の種類の楽器の単音ごとの前記 n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域を、前記第 1 の種類の楽器の単音に対応する前記第 2 の種類の楽器の前記単音の n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの特徴領域と置き換えることにより置換パワーエンベロープ・パラメータを作成するステップと、

前記楽器分類判定部が、前記第 1 の種類の楽器と前記第 2 の種類の楽器とが、同じ楽器分類に属すると判定したときには、前記倍音ピーク・パラメータ除く他のパラメータと前記置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成し、前記楽器分類判定部が、前記第 1 の種類の楽器と前記第 2 の種類の楽器とが、異なる楽器分類に属すると判定したときには、前記倍音ピーク・パラメータ及び前記パワーエンベロープ・パラメータを除く他のパラメータと前記置換倍音ピーク・パラメータ及び前記置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成するステップと、

前記合成分離音響信号と前記残差音響信号とを加算して、第 2 の種類の楽器から発生し

10

20

30

40

50



た楽器音を含む音楽音響信号を出力するステップとをコンピュータを用いて実施するために前記コンピュータで用いられる音楽音響信号生成用コンピュータプログラム。

【請求項 19】

請求項 16 乃至 18 のいずれか 1 項に記載の音楽音響信号生成用コンピュータプログラムが記録されたコンピュータ読み取り可能な記録媒体。

【請求項 20】

前記第 1 の種類の楽器または前記第 2 の種類の楽器を用いて演奏したときに前記第 1 の種類の楽器または前記第 2 の種類の楽器から発生する楽器音の音響信号を、前記分離音響信号分析保存部に保存された前記単音ごとの前記複数のパラメータを利用して生成するための操作を行う楽譜操作部を更に備えていることを特徴とする請求項 1 乃至 12 のいずれか 1 項に記載の音楽音響信号生成システム。

10

【請求項 21】

前記楽譜操作部は、前記他の楽譜の楽譜構造中の各単音にふさわしい、音高に関する音高パラメータ、音長に関する音長パラメータ及び調波モデルを構成するパラメータのうち音色に関わるパラメータを生成するように構成されている請求項 20 に記載の音楽音響信号生成システム。

【請求項 22】

演奏者がある楽譜を楽器で演奏して前記楽器から発生した楽器音の音響信号を含む音楽音響信号から抽出した、前記楽器音の音響信号のみを含む分離音響信号を単音ごとに保存する信号抽出保存部と、

20

前記単音ごとの分離音響信号を、少なくとも  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータと  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、前記単音ごとに前記複数のパラメータを分析して保存する分離音響信号分析保存部と、

前記楽譜とは異なる他の楽譜を前記演奏者が前記楽器を用いて演奏したときに前記楽器から発生する楽器音の音響信号を、前記分離音響信号分析保存部に保存された前記単音ごとの前記複数のパラメータを用いて生成するための操作を行う楽譜操作部とを含んでいることを特徴とする音楽音響信号生成システム。

【発明の詳細な説明】

30

【技術分野】

【0001】

本発明は、音楽音響信号の音色を変更することができる音楽音響信号生成システム及び方法並びに該方法をコンピュータで実施するために用いられるコンピュータプログラムに関するものである。

【背景技術】

【0002】

近年、音楽音響信号に特化し、楽器単位での音量の操作や音色の置換が可能な楽器音イコライザと呼ばれる新技術が開発されてきている。多くのオーディオプレーヤに実装されているイコライザは周波数帯域の操作によって楽曲の音響を変化させるが、楽器音イコライザが提供する楽器単位の操作によって音楽鑑賞の幅はさらに広がると期待される。非特許文献 1 に記載された吉井等の Drumix では、スネアドラムやバスドラムといった打楽器単位での音量操作と音色変更を実現している。一方、非特許文献 2 に示された糸山等の楽器音イコライザでは、打楽器だけではなく、全ての楽器単位の音量操作が可能だが、Drumix で実現されていた音色変更は扱われていない。なお非特許文献 2 に記載された発明を含むものとして、PCT/JP2008/57310 号 (W02008/133097) [特許文献 1] がある。

40

【先行技術文献】

【特許文献】

【0003】

50

【特許文献1】W O 2 0 0 8 / 1 3 3 0 9 7

【非特許文献】

【0004】

【非特許文献1】Yoshii, K., Goto, M. and G., O. H.: Drumix: An Audio Player with Realtime Drum-part Rearrangement Functions for Active Music Listening, IPSJ Journal, Vol. 48, No. 3, pp. 1229 ~ 1239 (2007)

【非特許文献2】糸山克寿, 後藤真孝, 駒谷和範, 尾形哲也, 奥乃博「楽譜情報を援用した多重奏音楽音響信号の音源分離と調波・非調波統合モデルの制約付パラメータ推定の同時実現」, 情報処理学会論文誌, Vol. 49, No. 3, pp. 1465 ~ 1479 (2008)

【非特許文献3】安部武宏, 糸山克寿, 吉井和佳, 駒谷和範, 尾形哲也, 奥乃博: 「音高による音色変化を考慮した楽器音の音高・音長操作手法」, 音楽情報科学研究会, Vol. 76, pp. 155 - 160 (2008).

10

【非特許文献4】Abe, T., Itoyama, K., Yoshii, K., Komatani, K., Ogata, T. and Okuno, H. G.: Analysis-and-Manipulation Approach to Pitch And Duration of Musical Instrument Sounds without Distorting Timbral Characteristics, International Conference on Digital Audio Effects, Vol. 11, pp. 249 - 256 (2008)

【非特許文献5】河原英紀「Vocoder のもう一つの可能性を探る - 音声分析変換合成システムSTRAIGHT の背景と展開」, 日本音響学会誌, Vol. 63, No. 8, pp. 442 - 449 (2007)

【非特許文献6】安部武宏, 糸山克寿, 吉井和佳, 駒谷和範, 尾形哲也, 奥乃博: 音色の音高依存性を考慮した楽器音の音高操作手法」, 情報処理学会論文誌, Vol. 50, No. 3 (2009)

20

【発明の概要】

【発明が解決しようとする課題】

【0005】

従来の技術では、任意の楽器パートをユーザの好みの音色に変更することはできなかった。また従来の技術では、未知の演奏の楽譜に対する演奏表情付き演奏音響信号を合成することはできなかった。

【0006】

本発明の目的は、既存の音楽音響信号中の任意の楽器パートの音色を任意の音色に変えることができる音楽音響信号生成システム及び方法並びに音色変更用コンピュータプログラムを提供することにある。

30

【0007】

本発明の他の目的は、既存の音楽音響信号中の任意の楽器パートの音色を用いて未知の演奏の楽譜に対する演奏表情付き演奏を合成することができる音楽音響信号生成システムを提供することにある。

【課題を解決するための手段】

【0008】

任意の楽器パートをユーザの好みの音色に変更することができれば、例えば、ロック風の楽曲を構成するギター、ベース、キーボードなどの楽器音を、ヴァイオリン、ウッドベース、ピアノなどの楽器音で置き換えることで、ユーザはその楽曲をクラシック風にアレンジして楽しむことができる。また、好きなギタリストが演奏した楽曲からギター音を抽出し、別の楽曲のギターパートをそのギター音で置き換えることで、ユーザはそのギタリストにさまざまなフレーズを演奏させることもできる。さらに、置き換える対象音から中間音を合成することにより、音色変更のバリエーションが広がると同時に幅の広い音楽鑑賞が可能となる。

40

【0009】

本願第1の発明の基本的な音楽音響信号の音色変更システムは、信号抽出保存部と、分離音響信号分析保存部と、置換用パラメータ保存部と、置換パラメータ作成保存部と、成分離音響信号生成部と、信号加算部とを備えている。

【0010】

50

信号抽出保存部は、第1の種類の楽器から発生した楽器音を含む音楽音響信号から抽出した、分離音響信号を単音ごとに保存し、また残差音響信号を保存する。分離音響信号は、第1の種類の楽器から発生した楽器音の単音のみを含む音響信号であり、残差音響信号は他の楽器の音響信号等のその他の音響信号を含むものである。音楽音響信号は、複数種類の楽器の音響信号が含まれる混合音響信号から分離されたものでも、最初から1つの楽器を演奏して得た単一楽器の音響信号でもよい。混合音響信号から音色変更の対象とする音響信号を分離するためには、公知の音響信号分離技術を実行する音響信号分離部を設ければよい。混合音響信号から音楽音響信号を分離する場合に、前述の非特許文献2に糸山等が提案した分離技術を用いると、他の楽器のパートの音響信号もすべて個別に分離することができ、同時に倍音ピーク・パラメータ等の各種のパラメータの分析もできる。

10

#### 【0011】

分離音響信号分析保存部は、単音ごとの分離音響信号を、少なくとも $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータ（通常、1単音あたり $n$ 個（ $n$ 次倍音分）の倍音ピーク・パラメータがある）と $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータ（通常、1単音あたり、倍音ピークの個数分のパワーエンベロープ・パラメータがある）によって定式化された調波モデルにより表現するために、1音あたりの複数のパラメータを分析して、分離音響信号中の複数音の単音についてそれぞれ複数のパラメータを保存する。このような複数のパラメータからなる調波モデルについては、前述の非特許文献2及びPCT/JP2008/57310号（W02008/133097：特許文献1）に詳しく説明されている。なお調波モデルは、少なくとも $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータと $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって構成されるものであればよく、特に上記非特許文献2に記載された調波モデルに限定されるものではない。例えば、調波モデルとして、倍音構造の非調和性を組み込んだ調波モデルを用いると、第1の種類の楽器が、弦楽器の場合におけるパラメータの生成精度を高めることができる。弦楽器音の倍音構造は厳密な整数倍をとらず、弦のスティフネスや長さによって各倍音ピークの周波数が若干高くなる。これは非調和性（インハーモニシティ）と呼ばれる。この非調和性は、周波数が高くなるほど影響が大きくなる。そこで非調和性を考慮した調波モデルを用いれば、第1の種類の楽器が弦楽器の場合に、倍音ピークの周波数の高い方向へのずれを考慮してパラメータを定めることができる。なおこの非調和性を考慮した調波モデルは、分析において利用するだけでなく、合成の際にも当然にして使用される。合成の際に調波モデルを使用する場合には、倍音構造の非調和性を示す変数（非調和度）は、音高依存特徴関数を用いて予測することができる。

20

30

#### 【0012】

1つの倍音ピーク・パラメータは、典型的には、周波数方向に現れる倍音ピークの強度を表す実数として表現される。またパワーエンベロープ・パラメータは、 $n$ 個の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータにそれぞれ含まれる同じ時間における倍音ピークのパワーの時間方向（周波数と同じで時間が異なる複数の倍音ピークのパワー）の変化を示すものであり、上記非特許文献2に記載されたパワーエンベロープ・パラメータに限定されるものではない。同じ楽器分類に属する楽器の音響信号であれば、各周波数におけるパワーエンベロープ・パラメータは、良く似た形状になる。例えば、ピアノ及び打弦楽器等の減衰楽器の単音のパワーエンベロープ・パラメータの形状は、大きく立ち上がった後に減衰する変化パターンを有している。またトランペット及び吹奏楽器等の持続楽器の単音のパワーエンベロープ・パラメータの形状は、立ち上がり部と立ち下がり部との間に緩やかな変化部分を持つ変化パターンを有している。なお保存する倍音ピーク・パラメータ及びパワーエンベロープ・パラメータのデータ形式は任意である。

40

#### 【0013】

置換用パラメータ保存部は、第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から作成した、音楽音響信号に含まれる全ての第1の種類の楽器の単音に対応する第2の種類の楽器から発生した複数の単音の音響信号を調波モデルにより表

50

現する場合に必要となる、第2の種類の楽器から発生した複数の単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータ及びパワーエンベロープ・パラメータを保存する。第2の種類の楽器から発生した複数の単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータは、事前に作成してもよい。作成した倍音ピーク・パラメータのデータ形式は、実数の形式でもよいが、関数の形式でもよく、任意である。また第2の種類の楽器から発生した楽器音の単音の音響信号は、信号抽出保存部に保存した全ての単音に対応するものを用意する必要はない。少なくとも第2の種類の楽器から発生した楽器音の音響信号として使用される単音信号が2単音分あれば、補間法等を用いてその他の単音の倍音ピーク・パラメータを作成してもよい。なお使用可能な単音の種類が多いほど、その他の単音の作成精度を高めることができるのは勿論である。

10

**【0014】**

置換パラメータ作成保存部は、分離音響信号分析保存部に保存された、第1の種類の楽器の単音ごとの $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、置換用パラメータデータ保存部に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成して保存する。置換倍音ピーク・パラメータは、全ての倍音ピーク・パラメータが第2の種類の楽器の楽器音から得られる倍音ピーク・パラメータによって置換されたものである。

20

**【0015】**

合成分離音響信号生成部は、分離音響信号分析保存部に保存された倍音ピーク・パラメータを除く他のパラメータと置換用パラメータ保存部に保存された置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。そして信号加算部は、合成分離音響信号と残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力する。

30

**【0016】**

本発明によれば、調波モデルを構成する複数のパラメータのうち音色に係わるパラメータの置換(変更)により、音色の変更(操作)をすることができるので、簡単に各種の楽器パートの音色変更を実現することができる。なお第1の種類の楽器の単音から得られるパワーエンベロープ・パラメータの変化パターンと第2の種類の楽器の単音から得られるパワーエンベロープ・パラメータの変化パターンとが近似していれば、音色の変更精度は高くなる。逆に両者の変化パターンが大きく異なると、音色は変わるものの、第2の種類の楽器の楽器音とは、第1の種類の楽器の雰囲気またはイメージが残った印象を受ける音色の変更となる。このような音色の変更も、ユーザによっては望まれる場合がある。音色の変更精度を高めるためには、パワーエンベロープ・パラメータの変化パターンが共通した楽器間の音色変更を実施するのが好ましい。

40

**【0017】**

そこで第2の発明では、置換用パラメータ保存部が、第2の種類の楽器の複数の単音ごとの $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータの他に、 $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータも保存する。そして置換パラメータ作成保存部は、置換倍音ピーク・パラメータを保存することに加えて、分離音響信号分析保存部に保存された、第1の種類の楽器の単音ごとの $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、置換用パラメータ保存部に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音についての、 $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータと置き換えることにより作成した置換パワーエンベロープ・パラメータを保存する。なおこの置き換えでは、時間方向の長さを一致させる必要があるときには、第2の種類の楽器のパワーエンベロープ・パラメータと音楽音響信号のパワーエンベロープ・パラメータのオンセットとオフセットが合うようにパワーエンベロープを伸縮させて置き換えを行う。なおこの音長操作に関しては、非特許文献3に記載されている。

50

## 【0018】

そして合成分離音響信号生成部は、分離音響信号分析保存部に保存された倍音ピーク・パラメータ及びパワーエンベロープ・パラメータを除く他のパラメータと置換パラメータ作成保存部に保存された置換倍音ピーク・パラメータ及び置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。その他は、第1の発明と同じである。このようにすると倍音ピークの置き換えだけでなく、第1の楽器の単音から得られるパワーエンベロープ・パラメータの変化パターンに変えて第2の種類の楽器の単音から得られるパワーエンベロープ・パラメータの変化パターンを用いることになるため、音色の変化の精度を高めることができる。

## 【0019】

第3の発明では、第2の発明の要件に加えて、第1の種類の楽器と第2の種類の楽器とが、同じ楽器分類に属するか否かを判定する楽器分類判定部をさらに備えている。第3の発明で用いる合成分離音響信号生成部は、楽器分類判定部が、第1の種類の楽器と第2の種類の楽器とが、同じ楽器分類に属すると判定したときには、第1の発明と同様に、分離音響信号分析保存部に保存された倍音ピーク・パラメータを除く他のパラメータと置換パラメータ作成保存部に保存された置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。また合成分離音響信号生成部は、楽器分類判定部が、第1の種類の楽器と第2の種類の楽器とが、異なる楽器分類に属すると判定したときには、分離音響信号分析保存部に保存された倍音ピーク・パラメータ及びパワーエンベロープ・パラメータを除く他のパラメータと置換パラメータ作成保存部に保存された置換倍音ピーク・パラメータ及び置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。このようにすると第2の種類の楽器がどのようなものであっても、自動的に最適な音色変更を行うことができる。

## 【0020】

また第3の発明では、楽器分類判定部を設けることに加えて、分離音響信号分析保存部が、単音ごとの分離音響信号中の非調波成分分布パラメータを分析して保存する機能を備えていてもよい。この場合、置換パラメータ作成保存部は、分離音響信号分析保存部に保存された、第1の種類の楽器の単音ごとの非調波成分分布パラメータを、置換用パラメータ保存部に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音の非調波成分分布パラメータと置き換えることにより作成した置換非調波成分分布パラメータ（第1の楽器の単音のオンセットにそそえた第2の楽器の単音の非調波成分分布パラメータ）を更に保存することになる。そして合成分離音響信号生成部は、分離音響信号分析保存部に保存された倍音ピーク・パラメータ、パワーエンベロープ・パラメータ及び非調波成分分布パラメータを除く他のパラメータと置換パラメータ作成保存部に保存された置換倍音ピーク・パラメータ、置換パワーエンベロープ・パラメータ及び置換非調波成分分布パラメータとを用いて、単音ごとの合成分離音響信号を生成するように構成することになる。このようにすると非調波成分も考慮して音色の変更を行えるので、更に音色の変更（操作）精度が高くなる。ただし、非調波成分分布パラメータは音色の操作に対する影響度が低いため、必ずしも考慮する必要はない。なお非調波成分分布パラメータの置換をするには、分離音響信号に調波成分だけでなく非調波成分も含む必要がある。そこで非調波成分分布パラメータを扱う場合には、上記非特許文献2に記載の調波モデル・非調波モデル統合モデルを用いることが必要になる。なお音楽音響信号が、混合音ではなく一種の楽器の単音だけから構成されている場合には、残差音響信号そのものを非調波成分とみなせるので、上記非特許文献2に記載の調波モデル・非調波モデル統合モデルを用いることなく、非調波成分分布パラメータの置換を適用することができる。

## 【0021】

なお置換用パラメータ保存部は、第2の種類の楽器から発生した楽器音の音響信号の前記複数種類の単音ごとの非調波成分分布パラメータを保存する機能を更に備えている。置換用パラメータ保存部は、パラメータ分析保存部とパラメータ補間生成保存部とを備えて構成してもよい。パラメータ分析保存部は、第2の種類の楽器から発生した楽器音の音響

10

20

30

40

50

信号から得た複数種類の単音ごとの分離音響信号を調波モデルにより表現する場合に必要な  
 となる、第2の種類の楽器から発生する複数種類の単音ごとの少なくともn次倍音成分の  
 相対強度を示す倍音ピーク・パラメータを分析して保存する。なお第2の種類の楽器から  
 発生する複数種類の単音についての、n次倍音成分の時間方向のパワーエンベロープを示  
 すパワーエンベロープ・パラメータは、事前に分析することにより得た倍音ピーク・パラ  
 メータと一緒にパラメータ分析保存部に保存してある。さらにパラメータ分析保存部には  
 、非調波成分分布パラメータを保存してある。そしてパラメータ補間生成保存部は、パラ  
 メータ分析保存部に保存した複数種類の単音についての倍音ピーク・パラメータに基づい  
 て、音楽音響信号に含まれる全ての単音に対応する第2の種類の楽器から発生する複数の  
 単音のうち複数種類の単音以外の単音についての音響信号をモデルにより表現する場合  
 必要となる第2の種類の楽器の複数の単音ごとの倍音ピーク・パラメータ及び前記パワー  
 エンベロープ・パラメータを補間法を用いて生成して保存する。このような構成を採用す  
 ると、第2の種類の楽器の単音のデータが少ない場合でも、置換に必要なパラメータを得  
 ることができる。なおパラメータ分析保存部は、分析により得られた前記n次倍音成分の  
 時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、代表パワーエ  
 ンベロープ・パラメータとして保存してもよい。

10

#### 【0022】

また置換用パラメータ保存部は、パラメータ分析保存部及びパラメータ補間生成保存部  
 に保存されたデータに基づいて、第2の種類の複数の単音ごとの倍音ピーク・パラメータ  
 を音高依存特徴関数として保存する関数生成保存部をさらに備えてもよい。この場合には  
 、置換パラメータ作成保存部は、第2の種類の楽器の単音の倍音ピーク・パラメータに含  
 まれる複数の倍音ピークを音高依存特徴関数から取得するように構成するのが好ましい。  
 このようにすると保存データのデータ量を少なくすることができる。また関数化すること  
 で複数の学習データの分析時の誤差を軽減することが期待される。

20

#### 【0023】

分離音響信号分析保存部が分析する複数のパラメータには、音高に関する音高パラメー  
 タと音長に関する音長パラメータ（なお音長パラメータはパワーエンベロープ・パラメー  
 タを包括する）とが含まれている場合には、音高パラメータを操作する音高操作部と、音  
 長パラメータを操作する音長パラメータ操作部をさらに備えているのが好ましい。これら  
 2つの操作部を備えていると、音色の変更（操作）の他に、音高、音長の変更（操作）も  
 行うことができる。

30

#### 【0024】

分離音響信号分析保存部が分析する複数のパラメータが第1の種類の楽器から発生した  
 すべての単音に対してそれぞれ別個に得られる場合には、楽譜構造と音響特徴の間の対応  
 をもとに、任意の構造の楽譜の各単音に対する音高パラメータ、音長パラメータ、及び音  
 色に関わるパラメータを構成するための楽譜操作部を備えることが可能である。

#### 【0025】

楽譜操作部は、「似た構造の楽譜は似たような音で演奏される」という仮定のもと、第1  
 の種類の楽器によって演奏された楽譜上の各単音に対応する音高パラメータ、音長パラメ  
 ータ、及び音色に関わるパラメータのすべてを用いて、ユーザが指定する任意の楽譜構造  
 中の各単音にふさわしい音高パラメータ、音長パラメータ、及び音色に関わるパラメータ  
 を生成する。ここでの「ふさわしさ」は、注目している単音の前後の単音との音高差など  
 から定義される。

40

#### 【0026】

そこで本発明の音楽音響信号生成システムでは、第1の種類の楽器または第2の種類の  
 楽器を用いて演奏したときに第1の種類の楽器または第2の種類の楽器から発生する楽器  
 音の音響信号を、分離音響信号分析保存部に保存された単音ごとの複数のパラメータを利用  
 して生成するための操作を行う楽譜操作部を更に備えていてもよい。楽譜操作部は、他  
 の楽譜の楽譜構造中の各単音にふさわしい、音高に関する音高パラメータ、音長に関する  
 音長パラメータ及び調波モデルを構成するパラメータのうち音色に関わるパラメータを生

50

成するように構成されている。

【 0 0 2 7 】

楽譜操作部の機能は音高操作部及び音長操作部を包含しているが、ユーザが指定する任意の楽譜構造が第 1 の種類の楽器によって演奏されていた楽譜と類似している場合、音高操作部及び音長操作部の機能によってユーザが指定する任意の楽譜構造中の各単音の音高パラメータ及び音長パラメータを変更した方が高い精度による操作が可能であるため、楽譜操作部の機能は必要に応じて音高操作部及び音長操作部の機能と使い分けることが望ましい。

【 図面の簡単な説明 】

【 0 0 2 8 】

【 図 1 】本発明の実施形態の音楽音響信号生成システムをコンピュータを用いて実現する場合の構成例を示すブロック図である。

【 図 2 】分離音響信号及び置換に用いる置換音響信号のパラメータ分析を説明するために用いる図である。

【 図 3 】 $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータを含む周波数エンベロープの一例を示す図である。

【 図 4 】 $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータ（時間エンベロープ）の一例を示す図である。

【 図 5 】本発明の実施の形態の一例としての、音楽音響信号生成システムの構成を示すブロック図である。

【 図 6 】周波数エンベロープの操作を示す図である。

【 図 7 】（ A ）乃至（ D ）は、トランペットの第 1 次倍音，第 4 次倍音，第 10 次倍音の相対強度、および調波成分と非調波成分のエネルギー比の音高特徴依存関数を示す図である。

【 図 8 】時間エンベロープの操作を説明するために用いる図である。

【 図 9 】音高軌跡の操作を説明するために用いる図である。

【 図 1 0 】（ A ）乃至（ C ）は、倍音ピーク間の相対強度、時間方向のパワーエンベロープ・パラメータ及び非調波成分の分布の例を示す図である。

【 図 1 1 】図 5 に示した実施の形態を具体的にコンピュータを用いて実現する場合に用いるコンピュータプログラムの一例のアルゴリズムを示すフローチャートである。

【 図 1 2 】置換用パラメータ保存部の具体的な構成を示す図である。

【 図 1 3 】音高依存特徴関数による置換パラメータの生成を説明するために用いる図である。

【 図 1 4 】倍音ピークの相対強度からのスペクトル包絡の導出を説明するために用いる図である。

【 図 1 5 】学習用特徴量を補間法を利用して生成する場合の式を説明するために用いる図である。

【 図 1 6 】同期パワーエンベロープ・パラメータ  $En(r)$  を得ることを説明するために用いる図である。

【 図 1 7 】パワーエンベロープ・パラメータの補間の概要図である。

【 図 1 8 】音楽音響信号内の単音のオンセットで同期をとることを示す図である。

【 図 1 9 】非調波成分分布パラメータの補間の概要図である。

【 図 2 0 】楽譜操作の概要を説明するために用いる図である。

【 図 2 1 】楽譜操作の概要を示す図である。

【 発明を実施するための形態 】

【 0 0 2 9 】

以下、本発明を実施するための形態（以下「実施形態」という。）を詳細に説明する。図 1 は、本発明の実施形態の音楽音響信号生成システムをコンピュータ 1 0 を用いて実現する場合の構成例を示すブロック図である。コンピュータ 1 0 は、CPU（Central Processing Unit）1 1 と、DRAM 等の RAM（Random Access Memory）1 2 と、ハードディスクドライブ（以下、「ハードディスク」という。）その他の大容量記憶手段 1 3 と、

10

20

30

40

50

フレキシブルディスクドライブ又はCD-ROMドライブ等の外部記憶部14と、LAN (Local Area Network) 又はインターネットなどの通信ネットワーク20との通信を行う通信部18とを備える。また、コンピュータ10は、キーボード又はマウス等の入力部15と、液晶ディスプレイ等の表示部16とを備えている。さらに、コンピュータ10は、MIDI音源等の音源17を搭載している。

#### 【0030】

CPU11は、パワースペクトルの分離処理と更新モデルパラメータのパラメータの推定(モデル適応)処理と、音色変更(操作)処理とを行うための各ステップを実行する計算手段として動作する。

#### 【0031】

音源17は、後述する入力音響信号を備えている。また音源分離を行うための入力音響信号と時間的に同期した標準MIDIファイル(Standard MIDI File、以下「SMF」という。)を楽譜情報データとして備えている。SMFはCD-ROM等や、通信ネットワーク20を介してハードディスク13に記録される。なお、「時間的に同期した」とは、SMF中の各楽器パートの単音(楽譜の音符に相当するもの)のオンセット時刻(発音時刻)と音長が実際の入力楽曲の音響信号における各楽器パートの単音と完全に同期しているということである。

#### 【0032】

なお、MIDI信号の記録、編集、再生などは、シーケンサ、またはシーケンスソフトウェア(図示を省略する)で行われる。ここでMIDI信号は、MIDIファイルとして扱われる。SMFとは、MIDI音源の演奏データを記録するための基本ファイルフォーマットである。SMFは、異なるシーケンサ、あるいはシーケンスソフトウェア間でのMIDIファイルの互換性を保つ統一規格である"チャンク"と呼ばれるデータ単位によって構成されている。SMFフォーマット化されたMIDIファイルデータのイベントとしては、大きく分けて、ミディ・イベント(MIDI Event)、システム・エクスクルーシブ・イベント(SysEx Event)、そしてメタ・イベント(Meta Event)の3種類がある。ミディ・イベントには、演奏データそのものが示されている。システム・エクスクルーシブ・イベントには、主にMIDIのシステム・エクスクルーシブ・メッセージが示されている。システム・エクスクルーシブ・メッセージは、特定の楽器にしかない情報を交換したり、特別の非音楽情報、イベント情報などを伝達するために用いられる。メタ・イベントには、テンポや拍子などの演奏全体の情報や、シーケンサやシーケンスソフトが利用する歌詞や著作権情報などの付加的な情報が示されている。全てのメタ・イベントは0xFFで始まり、次にイベント・タイプを表すバイトが続き、更にデータ長及びデータ自体が続く。MIDI演奏プログラムは、自分が認識できないメタ・イベントを無視するように設計されている。また、各イベントには、そのイベントを実行する時間的タイミングに関するタイミンダ情報が付加されている。このタイミング情報は、直前のイベントの実行からの時間差で示されている。例えば、このタイミング情報が"0"のときは、直前のイベントと同時にこのタイミング情報が付加されたイベントが実行されることになる。

#### 【0033】

一般に、MIDI規格を用いた音楽再生では、各種信号、楽器固有の音色をモデル化し、そのデータを格納した音源を各種パラメータで制御するシステムをとっている。そして、SMFの各トラックはそれぞれの楽器パートに対応しており、各楽器パートごとの分離信号が含まれている。また、SMFには、音高とオンセット時刻と音長又はオフセット時刻と楽器ラベルなどの情報が含まれている。

#### 【0034】

したがって、SMFが与えられていれば、これをMIDI音源で演奏することによって、入力される音響信号中の各単音にある程度近い音のサンプル(これを、「テンプレート音」という。)を生成することができる。テンプレート音から、ある楽器から生じる単音に対応した標準的なパワースペクトルにより表現されたデータのテンプレートを作成することができる。

10

20

30

40

50



## 【 0 0 3 5 】

テンプレート音またはテンプレートは、実際の入力音響信号の単音または単音のパワースペクトルとは完全には同一でなく、必ず音響的な違いがある。そのためテンプレート音またはテンプレートをそのまま分離音または分離のためのパワースペクトルとして使用することはできない。非特許文献2に糸山等が提案した音源分離システムを利用すれば、単音の更新パワースペクトルが後述する初期パワースペクトルに近い状態から、入力音響信号から分離した単音の最新のパワースペクトルに近い状態に徐々に近付ける学習（これを「モデル適応」という。）を行うことにより、更新モデルパラメータに含まれる複数のパラメータを最終的に望ましい形で収束させることができ、分離が可能になる。なお音源分離システムについては、その他の技術を用いることができるのは勿論である。

10

## 【 0 0 3 6 】

具体的な実施の形態の説明をする前に、本明細書中で使用する音色特徴を表現する音色特徴量を定義し、音楽音響信号（楽器音）の分析及び合成に用いる調波・非調波統合モデルについて説明する。

## 【 0 0 3 7 】

## [ 音色特徴量の定義 ]

ある楽器個体の実際の音がいくつか得られているとき、それらを元にして同個体の任意の音高・音長をもつ音、及び複数の音色特徴を含有する音を合成することにより合成音を得られる。このとき重要な点は、音色特徴が歪まないようにすることである。例えば、ある音高をもつ楽器音から音長操作により他の音高をもつ音を合成したとき、これらの音は同一の楽器個体から発せられていると感じられなければならない。

20

## 【 0 0 3 8 】

音色の音響特徴の歪みを抑えて楽器音を合成するために、以下の3つの特徴量を定義する。

## 【 0 0 3 9 】

(i) 倍音ピーク間の相対強度（倍音ピーク・パラメータ）

(ii) 非調波成分の分布（非調波成分分布パラメータ）

(iii) 時間方向エンベロープ（パワーエンベロープ・パラメータ）

音響心理学の分野では、音色の聴感上の知覚の差はおもに、(i) 高周波数領域での倍音ピークの有無、(ii) 発音時に発生する非調波成分、(iii) 各ピークの時間方向における振幅の変動、の3つに起因する傾向があると指摘されている。上記の音色特徴量は、これらの知見にそれぞれ対応する。

30

## 【 0 0 4 0 】

図2に、分離音響信号及び置換に用いる置換音響信号のパラメータ分析を説明するために用いる図である。前述の特徴量(i)及び(iii)は調波成分に関するもの、特徴量(ii)は非調波成分に関するものである。複数の実際の単音が与えられると、まず、各実際の単音の調波成分と非調波成分を分離後、各特徴量を分析することになる。

## 【 0 0 4 1 】

本実施の形態では、音色特徴量を分析するために、非特許文献2に示された糸山等が開発した調波・非調波統合モデルを拡張する。なお非特許文献2に示された調波・非調波統合モデルをそのまま利用してもよいのは勿論である。拡張した部分を以下に述べる。

40

## 【 0 0 4 2 】

## A．インハーモニシティ（非調和性）の組み込み

弦楽器音の倍音構造は厳密な整数倍をとらず、弦のスティフネスや長さによって各倍音ピークの周波数が若干高くなる。これは非調和性（インハーモニシティ）と呼ばれる。これを分析できるよう倍音ピークの周波数軸での配置間隔にインハーモニシティの理論式を適用した。

## 【 0 0 4 3 】

B．時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの実数表現  
ピアノ音やギター音といった急峻な立ち上がりを持つ楽器音のパワーエンベロープ・パ

50

ラメータを詳細に分析するために、ガウス関数の線形加算で表現されているパワーエンベロープ・パラメータを実数で表現した。

【0044】

本実施の形態では、上記拡張した調波・非調波統合モデルを用いて調波成分と非調波成分とを明示的に分けて取り扱う。すなわち、単音のスペクトログラム  $M(f, r)$  に対し、調波成分に対応するモデル  $M^{(H)}(f, r)$  と非調波成分に対応するモデル  $M^{(I)}(f, r)$  を  $(H)$  及び  $(I)$  で重み付けした混合モデルを次式のように表現する。

【数1】

$$M(f, r) = w^{(H)} M^{(H)}(f, r) + w^{(I)} M^{(I)}(f, r) \quad 10$$

【0045】

ここで、 $f$  と  $r$  はそれぞれパワースペクトルにおける周波数と時間を表す。また、 $\int_{f,r} M^{(I)}(f, r) df dr = 1$  という制約の付与により重み  $w^{(I)}$  は非調波成分のエネルギーと考えることができ、 $M^{(I)}(f, r)$  は非調波成分のスペクトログラムそのものを表す。一方、 $M^{(H)}(f, r)$  は、各倍音  $n$  に対するパラメトリックモデルの重み付き混合モデルとして表現される。

【数2】

$$M^{(H)}(f, r) = \sum_n F_n(f, r) E_n(r) \quad 20$$

【0046】

ここで、 $F_n(f, r)$  及び  $E_n(r)$  は、図3と図4に示すような  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータが一要素として含まれる周波数エンベロープ及び  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータ（パワーエンベロープ・パラメータ）を含むモデルとなっている。なお  $v_n$  が  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに対応する。なお非調波モデル  $M^{(I)}(f, r)$  が非調波成分分布パラメータに対応する。そして  $F_n(f, r)$  は混合正規分布を構成するひとつの要素の正規分布に混合比を乗じたものとして表現される。

【数3】

$$F_n(f, r) = \frac{v_n}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(f - \mu_n(r))^2}{2\sigma^2} \right] \quad 30$$

40

【0047】

ここで、 $\sigma$  は倍音ピークの周波数方向への分散、 $v_n$  は  $\sum_n v_n = 1$  を満たす重みであり、これが倍音ピーク・パラメータである。 $\mu_n(r)$  は  $n$  次倍音ピークの周波数軌跡で、次式のインハーモニシティの理論式に基づいて音高軌跡  $\mu(r)$  と非調和性を組み込むための非調和度  $B$  から次式のように表現される。

【数4】

$$\mu_n(r) = n\mu(r) \sqrt{1 + Bn^2}$$

50

## 【 0 0 4 8 】

ここで非調和性は弦楽器音の倍音ピーク特有の性質であり、弦の張力、硬さ、長さによって非調和度  $B$  が変わることになる。非調和性を有する倍音ピークの発生する周波数は上記式から求めることができる。注目する点は非調和度  $B$  を 0 にすれば  $\mu n(r) = n \mu(r)$  となり、非調和性の有無を非調和度  $B$  というパラメータで表現することができる。そこで非調和性を表現できるように調波モデルを拡張することで分析精度(モデル適応の精度)と合成時の音質(分析音の再現精度)の両方を高めることができる。そこで非調和性を表現できるように拡張された調波モデルを用いると、後述する分離音響信号分析保存部 3 及び置換用パラメータ保存部 4 においてより正確な倍音ピークの分析を提供することができる。なお基本的には、従来通りの調波モデル(非調和度  $B$  が 0 のモデル)を用いても、本発明の効果が得られることは勿論である。また非調和性は音高依存性を持っている。そのため音高の異なる楽器音(分離音響信号)の音高操作及び音色操作を行う場合には、後述する置換パラメータ作成保存部 6 において、音高依存特徴関数から予測した非調和度を用いるのが好ましい。 $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータ(パワーエンベロープ・パラメータ)  $E_n(r)$  は  $E_n(r)dr = 1$  を満たす関数である。この統合モデルにおいて、前述の音色特徴量(i), (ii) 及び(iii)は、それぞれ  $v_n$ ,  $M^{(1)}(f, r)$  及び  $E_n(r)$  (置換されるパラメータ)に対応する。これらの求め方は後に詳しく説明する。なおパワーエンベロープ・パラメータは、正弦波重畳モデルで扱われる振幅エンベロープとは異なり、各倍音ピークの時間方向のエネルギーの分布を表す。

10

20

## 【 0 0 4 9 】

## C. 楽器音の合成

調波成分に対応する調波信号  $s_H(t)$  を合成するには、特徴量(i) 及び(iii) をパラメータとする正弦波重畳モデルを用いる。非調波成分に対応する非調波信号  $s_I(t)$  を合成するには、特徴量(ii) を入力とするオーバーラップ加算法を用いる。各々に合成された調波信号と非調波信号を以下のように重ね合わせることで最終的な楽器音  $s(t)$  を合成する。

## 【 数 5 】

$$s(t) = s_H(t) + s_I(t)$$

30

## 【 0 0 5 0 】

ここで、 $t$  は信号のサンプル番地を表す。

## 【 0 0 5 1 】

図 5 は、上記に説明した拡張した調波・非調波統合モデルを用いる本発明の実施の形態の一例としての、音楽音響信号の音色変更システムの構成を示すブロック図である。この音楽音響信号の音色変更システムは、音響信号分離部 1 と、信号抽出保存部 2 と、分離音響信号分析保存部 3 と、置換パラメータ作成保存部 4 と、楽器分類判定部 5 と、置換用パラメータ保存部 6 と、合成分離音響信号生成部 7 と、信号加算部 8 と、音高操作部 9 A と、音長操作部 9 B とを備えている。

40

## 【 0 0 5 2 】

音響信号分離部 1 では、前述の拡張した調波・非調波統合モデルを用いて混合音楽音響信号から各音楽パートの音楽音響信号を分離する。調波・非調波統合モデルを用いる場合に、問題となるのは、先に示した統合モデルにおける未知パラメータ  $F_n(f, r)$ ,  $E_n(r)$ ,  $v_n$ ,  $\mu_n(r)$ ,  $M^{(1)}(f, r)$  を推定することである。そのため、非特許文献 2 の著者であり本願発明者の一人である糸山等は、統合モデルの単音のスペクトログラムとの Kullback-Leibler Divergence を減少させるようにパラメータを反復更新する手法を提案している。この反復過程は Expectation-Maximization アルゴリズムで、効率的にパ

50

ラメータを推定することができる。具体的には以下のコスト関数  $J$  を最小化することによって、本実施の形態で用いるモデルが単音のスペクトログラムへ適応される。

【数 6】

$$\begin{aligned}
 J = & \sum_n \iint \left( S_n^{(H)}(f, r) \log \frac{S_n^{(H)}(f, r)}{w^{(H)} E_n(r) F_n(f, r)} \right. \\
 & \left. - S_n^{(H)}(f, r) + w^{(H)} E_n(r) F_n(f, r) \right) df dr \\
 & + \iint \left( S^{(I)}(f, r) \log \frac{S^{(I)}(f, r)}{w^{(I)} M^{(I)}(f, r)} - S^{(I)}(f, r) + w^{(I)} M^{(I)}(f, r) \right) df dr \\
 & + \lambda^{(v)} \left( \sum_n v_n - 1 \right) + \sum_n \left( \lambda^{(E_n)} \left( \int E_n(r) dr - 1 \right) \right) \\
 & + \beta^{(I)} \iint \left( M^{(I)}(f, r) \log \frac{M^{(I)}(f, r)}{\bar{M}^{(I)}(f, r)} - M^{(I)}(f, r) + \bar{M}^{(I)}(f, r) \right) df dr \\
 & + \beta^{(E)} \int \left( \bar{E}(r) \log \frac{\bar{E}(r)}{E_n(r)} - \bar{E}(r) + E_n(r) \right) dr
 \end{aligned}$$

10

20

【0053】

ここで、 $M^{(I)}(f, r)$  は周波数方向で平滑化した非調波モデルである。非調波モデルは非常に高い自由度を持つため、調波モデルで表現すべき調波構造までも過剰に適応する。この非調波モデルの過剰適応を防ぐために、平滑化した非調波モデルとの距離をコスト関数に加算する。 $E_n(r)$  は各倍音ピークごとに平均したパワーエンベロープ・パラメータである。各倍音ピークのパワーが倍音ピーク間の相対強度とパワーエンベロープ・パラメータといったベクトル量及び調波エネルギーといったスカラー量の積算によって表現される。しかし、弱小なピークへのモデル適応時に、倍音ピーク間の相対強度が0に近くなるため、パワーエンベロープ・パラメータが非常に高い自由度を持つようになる。このため、音高操作時、音高依存特徴関数によって弱小だったピークの倍音ピーク間の相対強度が強くなったときに、高調波成分において強い歪みが発生する。この弱小ピークへのパワーエンベロープ・パラメータの過剰適応を防ぐため、平均化したパワーエンベロープ・パラメータとの距離をコスト関数に加算する。 $(v)$  と  $(E_n)$  はそれぞれ  $v_n$  と  $E_n(r)$  に対応するラグランジュの未定乗数項である。 $(I)$  と  $(E)$  はそれぞれ、非調波成分とパワーエンベロープ・パラメータへの制約重みとする。 $S_n^{(H)}(f, r)$  と  $S^{(I)}(f, r)$  はそれぞれ分離された各ピークの成分と非調波成分である。これらの分離は次式のように分配関数  $D_n^{(H)}(f, r)$  と  $D^{(I)}(f, r)$  をそれぞれ積算することによって行われる。

30

40

【数 7】

$$\begin{cases}
 S_n^{(H)}(f, r) = D_n^{(H)}(f, r) S(f, r) \\
 S^{(I)}(f, r) = D^{(I)}(f, r) S(f, r)
 \end{cases}$$

50

【 0 0 5 4 】

分離に用いられる分配関数はモデルのパラメータを固定してコスト関数  $J$  を最小化することによって求められ、次式のような方程式で導出される。

【 数 8 】

$$\begin{cases} D_n^{(H)}(f, r) = \frac{w^{(H)} E_n(r) F_n(f, r)}{w^{(H)} M^{(H)}(f, r) + w^{(I)} M^{(I)}(f, r)} \\ D^{(I)}(f, r) = \frac{w^{(I)} M^{(I)}(f, r)}{w^{(H)} M^{(H)}(f, r) + w^{(I)} M^{(I)}(f, r)} \end{cases} \quad 10$$

【 0 0 5 5 】

ただし、このとき最小化にあたって以下の制約を設ける。

【 数 9 】

$$\begin{cases} \sum_n D_n^{(H)}(f, r) + D^{(I)}(f, r) = 1 \\ 0 \leq D_n^{(H)}(f, r) \leq 1 \\ 0 \leq D^{(I)}(f, r) \leq 1 \end{cases} \quad 20$$

【 0 0 5 6 】

さらに、前述した非調波成分の自由度の高さを制限するために、次式のように非調波成分の分離に用いる分配関数に制約重み  $\gamma$  を積算する。

30

【 数 1 0 】

$$\begin{cases} D_n^{(H)}(f, r) = \frac{w^{(H)} E_n(r) F_n(f, r)}{w^{(H)} M^{(H)}(f, r) + \gamma w^{(I)} M^{(I)}(f, r)} \\ D^{(I)}(f, r) = \frac{\gamma w^{(I)} M^{(I)}(f, r)}{w^{(H)} M^{(H)}(f, r) + \gamma w^{(I)} M^{(I)}(f, r)} \end{cases} \quad 40$$

【 0 0 5 7 】

制約重み  $\gamma$  は反復過程の初期においては低い値が割り振られ、徐々に 1 に近づくように更新される。音響信号分離部 1 で、上記のモデルを用いて各楽器パートを構成する楽器音の音響信号の分離（分離音響信号の生成）と同時に、単音ごとの分離音響信号から上記パラメータを推定する。その結果、上記モデルを用いた場合には、音響信号分離部 1、信号抽出保存部 2 及び分離音響信号分析保存部 3 の大部分が実現される。上記モデルを用いない場合、音響信号分離部 1 では、公知の分離技術を用いて音楽音響信号を分離することになる。パラメータを推定することにより、1 つの音楽音響信号の分離が完了する。

50

## 【0058】

音響信号分離部1で音楽音響信号を分離した後、信号抽出保存部2は、音響信号分離部1において分離した第1の種類の楽器から発生した楽器音を含む音楽音響信号から抽出した、分離音響信号を単音ごとに保存し且つ残差音響信号を保存する。前述のように、非特許文献2の分離技術を用いる場合には、分離音響信号と残差音響信号の分離・抽出を行っている。なお音楽音響信号は、本実施の形態のように、音響信号分離部1を用いて複数種類の楽器の楽器音が含まれる混合音響信号から分離されたものでも、音響信号分離部1を用いずに、最初から1つの楽器を演奏して得た単一楽器の音楽音響信号でもよい。なお本実施の形態のように、混合音響信号から分離した音楽音響信号を用いる場合には、音響信号分離部1で分離した他の楽器のパートの音楽音響信号は、残差音響信号に含めることになる。

10

## 【0059】

分離音響信号分析保存部3は、単音ごとの分離音響信号を、少なくとも $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータ(通常、1単音あたり $n$ 次倍音分の $n$ 個の倍音ピーク・パラメータがある)と $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータ(通常、1単音あたり、倍音ピークの個数分のパワーエンベロープ・パラメータがある。)によって定式化された調波モデルにより表現するために、複数のパラメータを分析して保存する。なお非特許文献2に記載の調波・非調波統合モデルを音響信号分離部1で用いる場合には、分離音響信号分析保存部3は、音響信号分離部1に含まれることになる。なお調波モデルは、少なくとも $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータと $n$ 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって構成されるものであればよく、特に上記非特許文献2に記載された調波モデルに限定されるものではない。後に説明するように、調波モデルとして、倍音構造の非調和性を組み込んだ調波モデルを用いると、第1の種類の楽器が、弦楽器の場合におけるパラメータの生成精度を高めることができる。1つの倍音ピーク・パラメータは、典型的には、前述の図3に示すように、周波数方向に倍音ピークが並ぶパワースペクトル中の倍音ピークの強度の実数として表現される。図2のA欄には、第1の種類の楽器の楽器音の音響信号から作成したパラメータを示してある。図2において、A欄内の左端領域には、分析された $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータの一つを示している。そしてA欄の右端領域には、非調波成分のパワースペクトル(非調波成分分布パラメータ)が示されている。さらにA欄の中央領域には、分析された $n$ 次倍音成分の時間方向のパワーエンベロープ・パラメータの一つが示されている。パワーエンベロープ・パラメータは、図4に示すように、 $N$ 個の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータにそれぞれ含まれる同じ時間における倍音ピークのパワーの時間方向(周波数が同じで時間が異なる複数の倍音ピークのパワー)の変化を示すものであればよく、使用可能なパワーエンベロープ・パラメータは、上記非特許文献2に記載されたパワーエンベロープ・パラメータのみに限定されるものではない。

20

30

## 【0060】

置換用パラメータ保存部6は、第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から作成した、音楽音響信号に含まれる全ての単音に対応する第2の種類の楽器から発生する複数の単音についての音響信号を調波モデルにより表現する場合に必要な、第2の種類の楽器の複数の単音の $n$ 次倍音成分の相対強度を示す倍音ピーク・パラメータを保存する。また置換用パラメータ保存部6は、非調波成分分布パラメータも置換する場合には、第2の種類の楽器から発生した楽器音の音響信号の複数種類の単音ごとの非調波成分分布パラメータも保存する機能を備えている必要がある。

40

## 【0061】

図2のB欄には、第1の種類の楽器とは異なる第2の種類の楽器から発生した楽器音の音響信号から作成した、音楽音響信号に含まれる全ての単音に対応する第2の種類の楽器から発生する複数の単音についての音響信号を調波モデルにより表現する場合に必要な

50

る、第2の種類の楽器の単音のn次倍音成分の相対強度を示す倍音ピーク・パラメータの一例と、非調波成分と、n次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータの一例が示されている。

【0062】

同じ楽器分類に属する楽器の音響信号であれば、各周波数におけるパワーエンベロープ・パラメータは、良く似た形状になる。図1のA欄のパワーエンベロープ・パラメータの形状は、トランペット及び吹奏楽器等の持続楽器の単音のパワーエンベロープ・パラメータの形状であり、立ち上がり部と立ち下がり部との間に緩やかな変化部分を持つ変化パターンを有している。またB欄に示したパワーエンベロープ・パラメータの形状は、ピアノ及び打弦楽器等の減衰楽器の単音のパワーエンベロープ・パラメータの形状であり、大きな立ち上がりを持って減衰する変化パターンを有している。なお保存する倍音ピーク・パラメータ及びパワーエンベロープ・パラメータのデータ形式は任意である。非調波成分分布の形状も、楽器の形状に応じて相違する。非調波成分部分は、音の周波数をなす倍音ピーク以外の弱かな強度の周波数成分である。よって非調波成分分布パラメータも楽器の種類に応じて異なる。非調波成分分布の分析は、単音のみからなる音楽音響信号においては、考慮するのに十分に値する。

10

【0063】

第2の種類の楽器の複数の単音のn次倍音成分の相対強度を示す倍音ピーク・パラメータは、事前に作成してもよいが、本システムで作成してもよい。なお音響信号分離部1において混合音響信号から分離した他の楽器パートの音楽音響信号から得た単音を第2の種類の楽器音としても用いることができるのは勿論である。

20

【0064】

楽器分類判定部5は、第1の種類の楽器と第2の種類の楽器とが、同じ楽器分類に属するか否かを判定する。これは楽器分類が異なると、前述のパワーエンベロープのパターンが異なるためである。

【0065】

そして置換パラメータ作成保存部4は、分離音響信号分析保存部3に保存された、第1の種類の楽器の単音ごとのn次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、置換用パラメータデータ保存部6に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音のn次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成して保存する。置換倍音ピーク・パラメータは、全ての倍音パラメータが第2の種類の楽器の楽器音から得られる倍音パラメータによって置換されたものである。また置換パラメータ作成保存部4は、分離音響信号分析保存部3に保存された、第1の種類の楽器の単音ごとのn次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、置換用パラメータ保存部6に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音のn次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータと置き換えることにより作成した置換パワーエンベロープ・パラメータを保存する。なおこの置き換えでは、時間方向の長さを一致させる必要があるときには、第2の種類の楽器のパワーエンベロープ・パラメータと音楽音響信号のパワーエンベロープ・パラメータのオンセットとオフセットが合うようにパワーエンベロープを伸縮させて置き換えを行う。

30

40

【0066】

さらに置換パラメータ作成保存部4は、分離音響信号分析保存部3に保存された、第1の種類の楽器の単音ごとの非調波成分分布パラメータを、置換用パラメータ保存部に保存された、第1の種類の楽器の単音に対応する第2の種類の楽器の単音の非調波成分分布パラメータと置き換えることにより作成した置換非調波成分分布パラメータを更に保存する。

【0067】

合成分離音響信号生成部7は、楽器分類判定部5が、第1の種類の楽器と第2の種類の

50

楽器とが、同じ楽器分類に属すると判定したときには、分離音響信号分析保存部に保存された倍音ピーク・パラメータ除く他のパラメータと置換パラメータ作成保存部に保存された置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。また合成分離音響信号生成部 7 は、楽器分類判定部 5 が、第 1 の種類の楽器と第 2 の種類の楽器とが、異なる楽器分類に属すると判定したときには、分離音響信号分析保存部 3 に保存された倍音ピーク・パラメータ、パワーエンベロープ・パラメータ及び非調波成分分布パラメータを除く他のパラメータと置換パラメータ作成保存部に保存された置換倍音ピーク・パラメータ及び置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。このようにすると第 2 の種類の楽器がどのようなものであっても、自動的に最適な音色変更を行うことができる。そして信号加算部 8 は、合成分離音響信号生成部 7 から出力された合成分離音響信号と分離音響信号分析保存部 3 から得た残差音響信号とを加算して、第 2 の種類の楽器から発生した楽器音を含む音楽音響信号を出力する。図 2 の最下部には、残差音響信号を足す以前のパワースペクトルを示している。

10

#### 【0068】

本実施の形態によれば、調波モデルを構成するパラメータのうち音色に係わるパラメータの置換（変更）により、音色の変更（操作）をすることができるので、簡単に各種の音色変更を実現することができる。

#### 【0069】

なお楽器分類判定部 5 を設けず、また置換パラメータ作成保存部 4 には置換倍音ピーク・パラメータだけを保存するようにしてもよい。このようにすると、第 1 の種類の楽器の単音から得られるパワーエンベロープ・パラメータの変化パターンと第 2 の種類の楽器の単音から得られるパワーエンベロープ・パラメータの変化パターンとが近似していれば、音色の変更精度は高くなる。逆に両者の変化パターンが大きく異なると、所望の音色への変更精度は低くなるものの、第 2 の種類の楽器の楽器音とは、第 1 の種類の楽器の雰囲気またはイメージが残った印象を受ける音色の変更となる。このような音色の変更も、ユーザによっては望まれる場合があるので、許容される。

20

#### 【0070】

また置換の対象とするパラメータのうち、非調波成分分布パラメータの重要度は低いため、高い精度を要求されない場合には、置換の対象から除外してもよいのは勿論である。

#### 【0071】

本実施の形態では、分離音響信号分析保存部 3 が分析する複数のパラメータに、音高に関する音高パラメータと音長に関する音長パラメータとが含まれている。そこで音高パラメータを操作する音高操作部 9 A と、音長パラメータを操作する音長パラメータ操作部 9 B をさらに備えている。その結果、本実施の形態によれば、音高操作部 9 A 及び音長操作部 9 B を備えているので、音色の変更（操作）の他に、音高、音長の変更（操作）も行うことができる。

30

#### 【0072】

また本実施の形態では、分離音響信号分析保存部 3 が分析する複数のパラメータは第 1 の種類の楽器から発生したすべての単音に対してそれぞれ別個に得られる。そこでユーザが指定する任意の楽譜構造中の各単音にふさわしい、音高に関する音高パラメータ、音長に関する音長パラメータ、及び調波モデルを構成するパラメータのうち音色に関わるパラメータを生成するための楽譜操作部 9 C を備えている。本実施の形態では、楽譜操作部 9 C を備えているので、音色の変更（操作）の他に、楽譜の変更（操作）も行うことができる。

40

#### 【0073】

次に、音高、音長、音色、楽譜の操作（変更）の技術について説明する。JIS では、音色は「聴感上の音の性質の一つで、2 音の大きさ及び高さがともに等しくてもその 2 音が異なった感じを与えるとき、その相違に対応する性質」と定義されている。この定義では、音色は音高と音量とは独立の音の性質として扱われている。しかし、音色には音高への依存性があることが知られている。そのため音高によって変化するべき特徴量を保持した

50



まま音高操作を行うと操作された楽器音に音色の歪みが生じる。また音色に関係する物理量としてスペクトル包絡が知られている。しかし一つのスペクトル包絡だけで異なる音高の倍音ピーク間の相対強度を正確に表現することは出来ない。これら音色特徴量のみで音色の特徴を捉えられるとはいいがたい。そこで発明者は音色特徴量とそれらの依存関係を分析しなければ、音色の特徴を捉えることができないという立場で、音色特徴量に加え、複数の楽器音から音色特徴量の音高依存性を分析することで、楽器個体の音色を扱うことを試みた。すなわち、操作は音色特徴量の音高依存性を考慮して行う。そして最後に、調波成分・非調波成分を別々に再合成して、足し合わせる。

【0074】

発明者は、音高依存性を考慮した公知の論文[北原鉄朗, 後藤真孝, 奥乃博「音高による音色変化に着目した楽器音の音源同定:F0依存多次元正規分布に基づく識別手法」, 情報処理学会論文誌, Vol. 44, No. 10, pp. 2448.2458 (2003)]に着目した。この論文には、音高に対する音響特徴量を回帰関数(音高依存特徴関数)を用いて近似し、音高依存性を除去したあとの特徴量分布を学習することで、楽器音識別率が向上したと報告されている。なおこの論文中には、回帰関数を音高操作に使用することが開示されているだけで、音色置換にこの関数を用いることや、学習パラメータの補間生成を行うことは記載されていない。音色が音高に依存する理由として以下のことが知られている。

10

【0075】

音高を操作するには、音高軌跡 $\mu(r)$ に所望の倍率を乗算すればよいが、このとき音色特徴量の値を変化させずにそのまま利用することはできない。なぜなら、音色は音高依存性をもつことが知られており、音高の操作が大きくなるにつれて音色の歪みは増加するからである。

20

【0076】

図6に示すように、音高を $\mu(r)$ から $\mu'(r)$ に変化させる場合には、相対強度を $v_n$ から $v'_n$ へと適切に変化させる必要がある。

【0077】

この問題を解決するため、発明者は、北原鉄朗, 後藤真孝及び奥乃博が発表した「音高による音色変化に着目した楽器音の音源同定:F0依存多次元正規分布に基づく識別手法」[情報処理学会論文誌, Vol. 44, No. 10, pp. 2448.2458 (2003)]の論文で提案された音高依存性を考慮した楽器音識別手法に着目した。この論文では、音高に対する音響的特徴量を3次関数を用いて近似し、音高依存性を除去したあとの特徴量分布を学習することで、楽器音識別率が向上したと報告されている。

30

【0078】

音色が音高に依存する理由として以下が知られている。

【0079】

1. 音高が低くなれば、発音体は大きくなる。発音体の質量が大きくなると慣性も大きくなり、パワーエンベロープの立ち上がりや減衰により多くの時間を要する。

【0080】

2. 音高が高くなると振動損失が大きくなるために、高次の高調波は発生されにくくなる。

40

【0081】

3. 一部の楽器では音高により発音体が異なり、各発音体は異なる材質からできている。

【0082】

これらの知見から、楽器の音色は低域から高域にいくに従って連続的に変わるといえる。よって、本実施の形態では、音高よりも奏法に依存すると考えられる特徴量(iii)パワーエンベロープ・パラメータを除き、音高に対する特徴量(i)倍音ピーク間の相対強度(倍音ピーク・パラメータ), (ii)非調波成分の分布をn次関数(音高依存特徴関数と呼ぶ)(非調波成分分布パラメータ)で近似する。

【0083】

50

本実施の形態では、音高依存特徴関数の次数に3次を用いた。この次数は、限られた学習データから音色の音高依存性を学習でき、音色特徴量の音高による変化を十分に扱えるという基準を設け、予備実験より決定した。

#### 【0084】

具体的には、以下の2つのパラメータに着目した。

- (1) 各倍音の倍音ピーク間の相対強度  $v_n$
- (2) 調波成分のエネルギーに対する非調波成分のエネルギーの比  $(H)/(I)$

(1)の  $v_n$  に関しては、 $n$ 毎に独立に音高依存特徴関数を作成する。これによって、必ずしも  $v_n$  に関する制約  $\sum_n v_n = 1$  は満たされなくなるが、この場合で  $\sum_n v_n$  の値はほぼすべての音高に対して0.9~1.1程度に収まっており、生成される楽器音の音色がこれによって大きく変化することはないと考える。異なった音高をもつ複数のseed(単音)が与えられれば、それらの音色特徴量を分析し、最小二乗法によって音高依存特徴関数を求めることができる。得られた音高依存特徴関数を用いれば、所望の音高における音色特徴量を予測することができる。例として、図7(A)乃至(D)にトランペットの第1次倍音、第4次倍音、第10次倍音の相対強度、および調波成分と非調波成分のエネルギー比の音高特徴依存関数を示す。なお図7において、点と実線はそれぞれ、音高ごとに分析された音色の特徴量と、導出された音高依存特徴関数である。

#### 【0085】

音長を操作するには、パワーエンベロープ・パラメータ  $En(r)$  を所望の音長になるように伸縮させる方法は適切ではない。なぜなら、同一楽器個体では音長にかかわらず、発音の立ち上がり立ち下がり、および音高の変動周期は類似することが知られており、音長の操作が大きくなるにつれて歪みは増加するからである。特に楽器音の立ち上がり立ち下がりにはエネルギーが大きく変化する部分で音色の印象への関わりが深い。また、音高の変動周期は、特にビブラート奏法が多用される演奏される楽器にとっては重要であり、聴感に与える印象への影響が大きい。

#### 【0086】

この問題を解決するため、発明者はパワーエンベロープ・パラメータにおける立ち上がり立ち下がり部分を保存及び音高軌跡の時間的変動を再現する。まず、特徴量(iii)において、エネルギーの急峻な立ち上がり終了時をオンセット  $ron$ 、エネルギーの急峻な立ち下がり開始時をオフセット  $roff$  として定義する。音長を操作するには、図8に示すようにオンセット - オフセット区間のみを伸縮させればよい。また、図9に示すように、正弦波重畳モデルを用いてオンセット - オフセット区間の音高軌跡を表現し、操作前と同じ周波数特性をもつ所望の長さの音高軌跡を生成する。オンセット以前及びオフセット以降の音高軌跡は操作前のものを用い、オンセット - オフセット付近の軌跡はガウシアンによって平滑化される。

#### 【0087】

次に、楽譜の変更方法について説明する。本実施例において、楽譜を変更することは、変更後の楽譜中の各単音それぞれに対する音高軌跡、パワーエンベロープ・パラメータ、及び音色特徴量を用意することである。変更後の楽譜が変更前のものと本質的に異なる場合、前述の音高操作、音長操作によってこれらの特徴量を得ることは適切ではない。なぜなら、実演奏から分析した音高軌跡、パワーエンベロープ・パラメータ、及び音色特徴量には、楽譜構造に依存して発生する特徴量の揺らぎ、すなわち演奏表情が含まれているからである。従って、変更後の楽譜に対する上記特徴量は、変更前の楽譜演奏から得た特徴量を元に、「似た構造の楽譜は似たような音で演奏される」という仮定のもとで新たに生成するのが望ましい。

#### 【0088】

発明者は、図20に概念的に示すように、変更後の楽譜の全単音の特徴量を、1) 前音の音高、前音の音長、当該音の音高、当該音の音長の4要素が最も近い変更前の楽譜の単音と、2) 当該音の音高、当該音の音長、高音の音高、高音の音長の4要素が最も近い変更前の楽譜の単音、の2単音を分析して得た特徴量を混合比1:0 から0:1へと時間変化

10

20

30

40

50

させて重み付き混合する方法によって得る。この操作は変更前の楽譜演奏中で隣り合った音の組を、変更後の楽譜に合わせて次々と滑らかに連結させていく操作となる。

【0089】

次に、音色（変更）操作について説明する。音色を操作するには、各音色特徴量に実数の混合率を乗算する。各音色特徴量の補間方法には以下のような2通りがある。

【0090】

線形混合

【数11】

$$Feature^{(P)} = \sum_k \alpha_k Feature^{(k)}$$

10

【0091】

対数混合

【数12】

$$Feature^{(P)} = \exp \left[ \sum_k \alpha_k \log(Feature^{(k)}) \right]$$

20

【0092】

Feature には  $v_n$ ,  $M(l)(f, r)$ ,  $En(r)$  といった音色特徴量が当てはまる。また、 $k, P$  は、それぞれ各単音へのインデックスと、補間された特徴量へのインデックスである。各単音の混合率  $k$  は制約条件  $k \quad k = 1$  を満たし、 $0 < k < 1$  では内挿、 $1 < k$  あるいは  $k < 0$  では外挿となる。線形混合は、内挿と外挿の特徴量の変化率が一定であるが、音のエネルギーを対数的に捉える人間の聴覚特性を考慮できていない。これに対して対数混合は、人間の聴覚特性を考慮した補間方法であるが、混合した特徴量を最終的に指数化するため、外挿には注意が必要である。

30

【0093】

音色特徴量のアラインメントの取り方を図10に示す。図10(A)は、上段の第1の種類の楽器の単音ごとの  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと、第1の種類の楽器の単音に対応する下段の第2の種類の楽器の単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換える場合のアラインメントの取り方を示す。図10(B)は、第1の種類の楽器の単音から得られるパワーエンベロープ・パラメータと第2の種類の楽器の単音から得られるパワーエンベロープ・パラメータのアラインメントの取り方を示す。第2の種類の楽器のパワーエンベロープ・パラメータと第1の種類の楽器の単音のパワーエンベロープ・パラメータのオンセットとオフセットが合うようにパワーエンベロープを伸縮させて操作を行う。そして図10(C)は、上段の第1の種類の楽器の単音ごとの非調波成分と、下段の第2の種類の楽器の非調波成分のアラインメントの取り方を示す。両者のオンセットの部分を一致させるようにアラインメントを取ればよい。

40

【0094】

図11は、図5に示した実施の形態を具体的にコンピュータを用いて実現する場合に用いるコンピュータプログラムの一例のアルゴリズムを示すフローチャートである。また図13は、音色操作の状況を説明するために用いる図である。このプログラムでは、複数の単音ごとの  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータとパワーエンベロープ・パラメータを置換して、音色の変更（操作）を行う。まずステップST1では、第1の種類の楽器から発生した楽器音を含む音楽音響信号から、分離音響信号と残差音響信号と

50

を単音ごとにそれぞれ抽出する。またこのステップ S T 1 では、単音ごとの分離音響信号を、少なくとも n 次倍音成分の相対強度を示す倍音ピーク・パラメータと n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを含む複数のパラメータによって定式化された調波モデルにより表現するために、複数のパラメータを分析する（特徴量変換）。

#### 【 0 0 9 5 】

次にステップ S T 2 乃至ステップ S T 4 は、第 1 の種類の楽器とは異なる第 2 の種類の楽器から発生した楽器音の音響信号（置換音響信号）から、倍音ピーク強度及びパワーエンベロープに関する特徴量を抽出する。これらステップ S T 2 乃至ステップ S T 4 により、図 1 2 に示すように構成要素から構成される置換用パラメータ保存部 6 が構成される。すなわち図 1 2 に示す置換用パラメータ保存部 6 は、パラメータ分析保存部 6 1 と、パラメータ補間生成保存部 6 2 と、関数生成保存部 6 3 とを備えている。パラメータ分析保存部 6 1 は、ステップ S T 2 で実現される機能実現手段であり、第 2 の種類の楽器から発生した楽器音の音響信号から得た複数種類の単音の分離音響信号を調波モデルにより表現する場合に必要となる、複数種類の単音ごとの少なくとも n 次倍音成分の相対強度を示す倍音ピーク・パラメータ及び n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを分析して保存する。なおパラメータ分析保存部 6 1 は、分析により得られた n 次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、代表パワーエンベロープ・パラメータとして保存してもよい。

#### 【 0 0 9 6 】

図 1 3 の最上段には、置換音響信号の特徴量として、1 つの単音の n 個の n 次倍音成分の相対強度を示す倍音ピーク・パラメータのうちの 2 つの倍音ピーク・パラメータをパワースペクトルで示してある。パラメータ補間生成保存部 6 2 は、ステップ S T 3 で実現される機能実現手段である。ステップ S T 3 では、補間により学習用特徴量を生成する。具体的には、パラメータ分析保存部 6 1 に保存した複数種類の単音についての倍音ピーク・パラメータとパワーエンベロープ・パラメータとに基づいて、音楽音響信号に含まれる全ての単音に対応する第 2 の種類の楽器から発生する複数の単音のうち複数種類の単音以外の単音についての音響信号をモデルにより表現する場合に必要となる第 2 の種類の楽器の複数の単音ごとの倍音ピーク・パラメータ及びパワーエンベロープ・パラメータを補間法を用いて生成して保存する。このステップ S T 3 で行うことは、例えば 2 つの単音しかない場合において、必要なその他の複数の単音を補間法により生成して保存することである。

#### 【 0 0 9 7 】

ステップ S T 2 乃至ステップ S T 4 で、第 1 の種類の楽器とは異なる第 2 の種類の楽器から発生した楽器音の音響信号（置換音響信号）から倍音ピーク・パラメータ、パワーエンベロープ・パラメータ、非調波成分分布パラメータを抽出し、これらを補間することで置換に用いる各々のパラメータ（置換パラメータ）を生成する。補間により置換パラメータを生成することで、音色置換を所望する音楽音響信号中の単音と同じ音高、音長を持つ第 2 の種類の楽器の音響信号を限られた数の置換音響信号で置換することができる。音色は音高依存性を持っており、特に、倍音ピーク・パラメータは特に強い音高依存性を持っていることが非特許文献 4 の実験で知られている。

#### 【 0 0 9 8 】

一方、スペクトル包絡は小さな音高依存性しか持っておらず、非特許文献 5 でスペクトル包絡を保持した品質の高い音声の音高操作手法が報告されている。

#### 【 0 0 9 9 】

このスペクトル包絡を保持した音高操作手法は、非特許文献 4 にて評価実験における比較対象ともなっており、実験でスペクトル包絡の音高依存性の小ささが示されている。また、音響心理学の分野では、各ピークの時間方向における振幅の変動と発音時に発生する非調波成分で音色の時間変化が知覚される傾向があると指摘されている。すなわち、パワーエンベロープ・パラメータは発音時及び発音が持続している間、非調波成分分布パラメ

ータは発音時が音色の知覚に重要な特徴を含んでいるといえる。

【0100】

本実施の形態の倍音ピーク・パラメータの補間では、倍音ピーク・パラメータよりもスペクトル包絡のほうが、音高依存性が小さいことに着目し、倍音ピーク・パラメータをスペクトル包絡に変換する。スペクトル包絡 $v(f)$ への変換は、図14に示すように隣接する各倍音ピーク・パラメータ $v_n$ を補間(線形補間、スプライン補間etc.)することで実現される。このとき、補間区間を超えた周波数(音高より下及び最高次の倍音ピーク周波数より上)のスペクトル包絡の変換には最近傍周波数の倍音ピーク・パラメータを用いる。以下、同様に補間区間を超えた範囲の補間には最近傍に位置するパラメータ値を用いるものとする。

10

【0101】

さらに、変換して得られたスペクトル包絡 $v(f)$ を次式に基づいて補間することで、音色置換を所望する音楽音響信号中の任意の音高 $\mu$ を持つ単音の補間スペクトル包絡を得られる。

【数13】

$$\hat{v}(f) = \exp \left[ (1 - \alpha) \log(\bar{v}^{(k)}(f)) + \alpha \log(\bar{v}^{(k+1)}(f)) \right]$$

20

【0102】

ここで、 $k$ は置換音響信号に付与されているインデックスであり、 $v^{(k)}(f)$ 、 $v^{(k+1)}(f)$ はそれぞれ低域、高域において最近傍の音高を持つ置換音響信号のスペクトル包絡である。また、 $\alpha$ はこれら置換音響信号の音高 $\mu^{(k)}$ 、 $\mu^{(k+1)}$ から決定される補間率であり、次式により決定される。

【数14】

$$\alpha = \frac{\hat{\mu} - \mu^{(k)}}{\mu^{(k+1)} - \mu^{(k)}}$$

30

【0103】

ここで音高 $\mu_n$ は以下のように定義される。

【数15】

$$\mu_n = n\mu\sqrt{1 + Bn^2}$$

40

【0104】

最後に、次式のように各倍音ピーク周波数の補間スペクトル包絡から補間倍音ピーク・パラメータを得る。

【数16】

$$\hat{v}_n = \hat{v}(\hat{\mu}_n)$$

【0105】

50

以上の倍音ピーク・パラメータの補間の概要図を図 15 に示す。

【0106】

本実施の形態のパワーエンベロープ・パラメータの補間では、発音時及び発音の持続している間が各ピークの振幅における音色の知覚に重要であることに着目し、置換音響信号のオンセットとオフセットを、置換を所望する音楽音響信号内の単音のオンセットとオフセットに同期させる。同期させるオンセット  $ron$  とオフセット  $roff$  はそれぞれ、平均パワーエンベロープ・パラメータにおけるパワーが十分に大きくなった地点とパワーが急峻に減少する地点を表すものであり、検出には手法を問わない。置換を所望する音楽音響信号内の単音のオンセット  $ron$ 、オフセット  $roff$  で同期をとるにはパワーエンベロープ・パラメータを時間軸上で操作する必要がある。これには非特許文献 6 で報告されている手法を用い、図 16 に示すようにオンセット・オフセット区間 ( $ron - roff$ ) のみを操作することで同期パワーエンベロープ・パラメータ  $En(r)$  を得る。

10

【0107】

同期パワーエンベロープ・パラメータ  $En(r)$  を次式に基づいて補間することで、音色置換を所望する音楽音響信号中の任意の音長を持つ単音の補間パワーエンベロープ・パラメータ  $En(r)$  を得られる。

【数 17】

$$\hat{E}_n(r) = \exp \left[ (1 - \alpha) \log(\bar{E}_n^{(k+1)}(r)) + \alpha \log(\bar{E}_n^{(k)}(r)) \right]$$

20

【0108】

ここで、 $E^{(k)n}(f)$ 、 $E^{(k+1)n}(f)$  はそれぞれ低域、高域において最近傍の音高を持つ置換音響信号のパワーエンベロープ・パラメータである。パワーエンベロープ・パラメータの補間にも倍音ピーク・パラメータの補間で用いた補間率も用いる。以上のパワーエンベロープ・パラメータの補間の概要図を図 17 に示す。

【0109】

本実施の形態の非調波成分分布パラメータの補間では、発音時が非調波成分における音色の知覚に重要であることに着目し、置換音響信号のオンセットを置換を所望する音楽音響信号内の単音のオンセットに同期させる。同期させるオンセット  $ron$  はパワーエンベロープ・パラメータの同期で用いたものと同様である。置換を所望する音楽音響信号内の単音のオンセット  $ron$  で同期をとるには、図 18 に示すように非調波成分分布パラメータを時間軸上で平行移動させればよく、これにより同期非調波成分分布パラメータ  $M^{(l,k)}(f, r)$  を得る。同期非調波成分分布パラメータ  $M^{(l,k)}(f, r)$  を次式に基づいて補間することで、音色置換を所望する音楽音響信号中の任意の音長を持つ単音の補間非調波成分分布パラメータ  $M^{(l,k)}(f, r)$  を得ることができる。

30

【数 18】

$$\hat{M}^{(l,k)}(f, r) = \exp \left[ (1 - \alpha) \log(\bar{M}^{(l,k)}(f, r)) + \alpha \log(\hat{M}^{(l,k+1)}(f, r)) \right]$$

40

【0110】

ここで、 $M^{(l,k)}(f, r)$ 、 $M^{(l,k+1)}(f, r)$  はそれぞれ低域、高域において最近傍の音高を持つ置換音響信号の非調波成分分布パラメータである。非調波成分分布パラメータの補間にも倍音ピーク・パラメータの補間で用いた補間率も用いる。以上の非調波成分分布パラメータの補間の概要図を図 19 に示す。さらに、倍音ピーク・パラメータ及び非調波成分分布パラメータを構成する非調波成分エネルギー  $w(l)$  においては、関数化することで置換音響信号のパラメータ分析時の誤差を軽減させることができる。このとき、補間に用

50

いる置換音響信号は多いほうが望ましい。この関数化には非特許文献 5 で報告されている音高依存特徴関数を用い、倍音ピーク・パラメータ及び非調波成分分布パラメータを学習した音高依存特徴関数から予測する。

【0111】

ステップ S T 4 では、音高依存性特徴関数の学習を行う。なお学習方法及び学習させるパラメータは、前述の音高操作時に使われる音高依存性特徴関数と同じである。ステップ S T 4 により、図 1 2 の関数生成保存部 6 3 が構成されている。関数生成保存部 6 3 は、パラメータ分析保存部 6 1 及びパラメータ補間生成保存部 6 2 に保存されたデータに基づいて、第 2 の種類の複数の単音ごとの倍音ピーク・パラメータを音高依存特徴関数として保存する。具体的にステップ S T 4 では、ステップ S T 3 で生成したいくつかの単楽器音の特徴量から最小二乗法によって回帰関数の係数を推定する（図 1 3 の上から 3 段目の図参照）。この回帰関数は、音高依存特徴関数と呼ばれるものである。具体的には、1 つの単音についての  $n$  個の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータから、同じ周波数を持って発生する倍音ピークを各次元のデータから（1 次から  $n$  次まで）集めてそれらの包絡線を表したものである。このような関数が得られれば、第 2 の種類の楽器の単音の倍音ピーク・パラメータに含まれる複数の倍音ピークを各次元の音高依存特徴関数から取得することができる。このように関数化することで、複数の学習データの分析時の誤差を軽減することができる。

10

【0112】

本発明において、ステップ S T 4 を用いた関数化は必須要件ではなく、ステップ S T 3 の精度が高ければ、ステップ S T 3 で取得したデータをそのまま利用してもよい。また第 2 の種類の楽器の複数の単音ごとの必要なパラメータは、どのように作成してもよく、本実施の形態に限定されるものではない。

20

【0113】

図 1 1 に戻って、ステップ S T 5 では、第 1 の種類の楽器の単音ごとの  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークを、第 1 の種類の楽器の単音に対応する第 2 の種類の楽器の単音の  $n$  次倍音成分の相対強度を示す倍音ピーク・パラメータに含まれる複数の倍音ピークと置き換えることにより置換倍音ピーク・パラメータを作成する。ステップ 5 では、ステップ S T 4 で求めた音高依存特徴関数から置き換えに必要な第 2 の楽器の倍音ピークを取得している。そしてステップ S T 6 では、第 1 の種類の楽器と第 2 の種類の楽器とが、同じ楽器分類に属するか否かを判定する。ステップ S T 6 で、第 1 の種類の楽器と第 2 の種類の楽器とが、同じ楽器分類に属すると判定したときには、ステップ S T 8 へと進む。ステップ S T 6 で、第 1 の種類の楽器と第 2 の種類の楽器とが、同じ楽器分類に属するものではないと判定したときには、ステップ S T 7 へと進む。ステップ S T 7 では、ステップ S T 2 乃至ステップ S T 4 で求めた、第 2 の種類の楽器の複数の単音の  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを取得する。そして第 1 の種類の楽器の単音ごとの  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータを、第 1 の種類の楽器の単音に対応する第 2 の種類の楽器の単音の  $n$  次倍音成分の時間方向のパワーエンベロープを示すパワーエンベロープ・パラメータと置き換えることにより置換パワーエンベロープ・パラメータを作成する。非調波成分分布パラメータについても、このステップ S T 7 で置換非調波成分分布パラメータを作成する。

30

40

【0114】

ステップ S T 6 で、2 つの楽器が同じ楽器分類に属することが判定された場合、ステップ S T 8 では、分離音響信号分析保存部に保存された倍音ピーク・パラメータ除く他のパラメータと置換用パラメータ保存部に保存された置換倍音ピーク・パラメータとを用いて、単音ごとの合成分離音響信号を生成する。ステップ S T 6 で、2 つの楽器が同じ楽器分類に属するものではないことが判定された場合、ステップ S T 8 では、倍音ピーク・パラメータ及びパワーエンベロープ・パラメータを除く他のパラメータと置換倍音ピーク・パラメータ及び置換パワーエンベロープ・パラメータとを用いて、単音ごとの合成分離音響

50

信号を生成する。そして最後のステップST9では、単音ごとの合成分離音響信号と残差音響信号とを加算して、第2の種類の楽器から発生した楽器音を含む音楽音響信号を出力する。

【0115】

図11のアルゴリズムでは、ステップST6で、楽器分類を判定しているが、楽器分類の判定はステップST5よりも前において行ってもよい。また当初より、同じ楽器分類に属する楽器の音響信号間でのみ音色変更をすることが決まっている場合には、ステップST7は不要であり、ステップST2乃至ステップST4でもパワーエンベロープ・パラメータを扱う必要はない。

【0116】

次に、図1の実施の形態を処理する具体的な実装について説明する。

【0117】

[音高操作]

音高操作を行うには、周波数エンベロープを構成する音高軌跡 $\mu(r)$ に対して、実数 $\alpha$ （音高を低くする場合： $0 < \alpha < 1$ 、音高を高くする場合： $1 < \alpha$ ）を乗算する。ここで、 $\mu(r)$ を所望する操作後の音高とすると以下が成り立つ。

【数19】

$$\hat{\mu}(r) = \alpha\mu(r)$$

【0118】

例えば、 $\alpha$ を2とすれば、seed（単音）の1オクターブ上の音高の楽器音が合成できる。操作後の楽器音の倍音ピーク間の相対強度 $v_n$ は、音高依存特徴関数から予測される各倍音ごとの倍音ピーク間の相対強度を制約条件 $\sum_n v_n = 1$ より正規化することで得られる。また、操作後の楽器音の非調波成分のエネルギー $E_{non}^{(H)}$ は、調波成分のエネルギー $E_{har}^{(H)}$ を音高特徴依存関数から予測される調波成分に対する非調波成分のエネルギーの比 $\kappa$ で割ることで得られる。

【0119】

[音長操作]

音長操作を行うには、オンセット・オフセット間の時間方向エンベロープ $E_n(r)$ と音高軌跡 $\mu(r)$ を操作する。操作によって得られた時間方向エンベロープと音高軌跡をそれぞれ $E_n$ と $\mu(r)$ とする。

【0120】

[オンセットとオフセットの検出]

本願明細書におけるオンセットとは、楽器音の時間方向の振幅が十分に大きくなってから、振幅の変動が一定になる瞬間である。オフセットとは、時間方向の振幅が十分な大きさを持っており、振幅の変動が一定の状態が得られなくなる瞬間である。この定義に従い、オンセットとオフセットを以下の通り検出する。

【数20】

$$r_{on} = \min \left( r \mid \text{abs} \left[ \frac{d\bar{E}(r)}{dr} \right] \leq \epsilon, \bar{E}(r) \geq \kappa \right)$$

$$r_{off} = \max \left( r \mid \text{abs} \left[ \frac{d\bar{E}(r)}{dr} \right] \leq \epsilon, \bar{E}(r) \geq \kappa \right)$$

【0121】

10

20

30

40

50



ここでThは、楽器音の時間方向の振幅の十分な大きさを示す閾値である。持続系の楽器はこれでもよいが、打弦楽器や撥弦楽器といった減衰楽器のオンセットとオフセットは、ほぼ同時時刻となり、オンセット・オフセット間を伸縮させることができない。よって、シンセサイザーにおける減衰楽器の振幅制御を参考に、パワーエンベロープ・パラメータの終端を減衰楽器音のオフセットとみなし、オンセット以降のパワーエンベロープ・パラメータを伸縮の対象とする。

【0122】

[楽譜操作]

ユーザが指定する変更後の楽譜の各単音の特徴量は、分析した変更前（元演奏）の楽譜との楽譜構造の類似性に基づいて生成される。図21は、楽譜操作における操作の流れを示しており、変更前の楽譜演奏音響信号から演奏表情を含む特徴量を抽出し、これを用いて楽譜構造の類似性に基づき変更後の楽譜に対する特徴量を生成する。そこで発明者は、変更後の楽譜の第j音に対する特徴量Featureを、変更前の楽譜中のノートナンバーNと音長Lの類似する単音の特徴量から算出する方法をとった。まず、変更後の楽譜の第j音に対して以下の条件を満たす分析済変更前の楽譜中の2音を選出する。

【数21】

$$q_j^- = \operatorname{argmin}_k \sum_{p=-1,0} (|\bar{N}_{j+p} - N_{k+p}| + \alpha |\bar{L}_{j+p} - L_{k+p}|)$$

$$q_j^+ = \operatorname{argmin}_k \sum_{p=0,1} (|\bar{N}_{j+p} - N_{k+p}| + \alpha |\bar{L}_{j+p} - L_{k+p}|)$$

【0123】

ここで、 $N_k$ ,  $L_k$  は変更前の楽譜のノートナンバーと音長であり、 $N_j$ ,  $L_j$  は変更後の楽譜のノートナンバーと音長であり、 $\alpha$  はそれらの重みを決定する定数である。次に、得られた二つの単音の特徴量を混合して、第j音にふさわしい音モデルを算出する。

【数22】

$$Feature^{(j)}(r) = \left(1 - \frac{r}{\bar{L}_j}\right) Feature^{(q_j^-)}(r) + \frac{r}{\bar{L}_j} Feature^{(q_j^+)}(r)$$

【0124】

ただし、上記式において、 $Feature^{(j)}(r)$  は第j音の特徴量中の時間フレームrに対するものであり、その四則演算は各パラメータ同士のもので定義する。また、

【外1】

【0125】

$$Feature^{(q_j^-)}(r), Feature^{(q_j^+)}(r)$$

はそれぞれ変更前の楽譜の $q_j^-$ 音、 $q_j^+$ 音の特徴量を音高が $N_j$ 、音長が $L_j$ となるように操作をしたものである。この式は二つの音特徴量の混合比を1:0から0:1へと

10

20

30

40

50

時間変化させることを意味しており、 $q_j^+ = q_{j+1}$  であることから、変更前の楽譜中で隣り合った音の組を変更後の楽譜の楽譜に合わせて次々と滑らかに連結させていく操作となっている。

【 0 1 2 6 】

[ 音高軌跡のモデル化 ]

オンセット・オフセット間の音高軌跡  $\mu(r)$  をモデル化するため、音高の周期的変動が時不変であることを仮定し、正弦波重畳モデルに基づく音高軌跡モデルを構築する。すなわち、音長操作後の音高軌跡は次式のように表現される。

【 数 2 3 】

$$\hat{\mu}(r) = \sum_k A_k^{(\mu)} \exp[j\omega_k^{(\mu)} r + \phi_k^{(\mu)}] + \int \mu(r) dr / R$$

10

【 0 1 2 7 】

ここで、 $R$  はフレーム数である。未知パラメータは、音高軌跡を構成する各正弦波の振幅  $A_k(\mu)$  と、周波数  $k(\mu)$ 、位相  $\phi_k(\mu)$  である。これらは既存の正弦波重畳モデルのパラメータ推定手法によって導出することが可能である。

【 0 1 2 8 】

20

[ 音色操作 ]

補間された各音色特徴量は次式によって得られる。

【 数 2 4 】

$$Feature^{(P)} = \exp \left[ \sum_k \alpha_k \log(Feature^{(k)}) \right]$$

30

【 0 1 2 9 】

ここで  $Feature$  には、 $v_n, M^{(1)}(f, r), E_n(r)$  といった音色特徴量が当てはまる。また、 $k, P$  は、それぞれ各  $seed$  (単音) へのインデックスと、補間された特徴量へのインデックスである。倍音ピーク間の相対強度  $v_n$  ではアラインメントは不要である。非調波成分の分布  $M^{(1)}(f, r)$  ではオンセットのみでアラインメントがとられる。一方、時間方向の振幅エンベロープ  $E_n(r)$  ではオンセットとオフセットが揃うように音長操作された後アラインメントがとられる。

【 0 1 3 0 】

[ 楽器音の合成 ]

調波モデルから調波信号  $s_H(t)$  を、非調波モデル  $s_I(t)$  から非調波信号を合成し、以下のように重ね合わせることで最終的な楽器音  $s(t)$  を合成する。

40

【 数 2 5 】

$$s(t) = s_H(t) + s_I(t)$$

【 0 1 3 1 】

ここで、 $t$  はサンプリングされた信号のサンプル番地を表す。

【 0 1 3 2 】

[ 調波信号の合成 ]

50

調波信号 $s_H(t)$ を合成するには、次式によって表現される正弦波重畳モデルを用いる。

【数 2 6】

$$s_H(t) = \sum_n A_n(t) \exp[j\phi_n(t)]$$

【0 1 3 3】

ここで、 $A_n(t)$ 、 $\phi_n(t)$ とはそれぞれ $n$ 番目の正弦波の瞬時振幅と瞬時位相である。このモデルでは、各正弦波の振幅と周波数が定常性を持っていることが仮定されている。瞬時位相は、フレーム単位で分析されている音高軌跡をスプライン補間によってサンプル単位を補間した操作後の音高軌跡 $\mu(t)$ を積分することによって得られる。

10

【数 2 7】

$$\phi_n(t) = \phi_n(0) + n\sqrt{1 + Bn^2} \int_0^t \hat{\mu}(\tau) d\tau$$

【0 1 3 4】

ここで、 $\phi_n(0)$ は任意の初期位相である。正弦波重畳モデルではトラッキングしたピークを瞬時振幅として用いる。調波構造の概形をモデル化した調波モデルにおいては、周波数エンベロープを構成する各ガウス関数の平均にパワーエンベロープ・パラメータと調波エネルギーを積算したものをトラッキングしたピークとみなすことができる。特徴量抽出のモデルと楽器音合成のモデルが異なるために合成音を持つ倍音の相対強度は分析対象の楽器音のものとは必ずしも一致しないが、実験的にはこの操作を経ても特徴量が大きく変化することはなかったため、モデルの違いの音色への影響は小さいと考える。よって、瞬時振幅は次式から求めることができる。

20

【数 2 8】

$$A_n(t) = \frac{w^{(H)} \hat{v}_n \hat{E}_n(t)}{\sqrt{2\pi\sigma}}$$

30

【0 1 3 5】

ここで、時間方向エンベロープ $E_n(r)$ にはスプライン補間を用いてサンプル単位にしたものが用いられる。

【0 1 3 6】

[非調波信号の合成]

非調波信号 $s_l(t)$ を合成するには、オーバーラップ加算法を用いる。このとき、非調波エネルギー $(l)$ を乗算した非調波モデル $(l)M^{(l)}(f, r)$ をスペクトログラムとみなして信号に変換する。位相はseedのものをそのまま利用する。

40

【0 1 3 7】

次に、オンセット・オフセット情報に基づく制約を付加したコスト関数の使用について説明する。

【0 1 3 8】

以下に示すコスト関数を最小化することによって、調波非調波統合モデルを分離の対象音が存在する混合音へ適応する。

【数 2 9】

$$\begin{aligned}
J = & \sum_n \iint \left( S_n^{(H)}(f, r) \log \frac{S_n^{(H)}(f, r)}{w^{(H)} E_n(r) F_n(f, r)} - S_n^{(H)}(f, r) + w^{(H)} E_n(r) F_n(f, r) \right) df dr \\
& + \iint \left( S^{(I)}(f, r) \log \frac{S^{(I)}(f, r)}{w^{(I)} M^{(I)}(f, r)} - S^{(I)}(f, r) + w^{(I)} M^{(I)}(f, r) \right) df dr \\
& + \lambda^{(v)} \left( \sum_n v_n - 1 \right) + \sum_n \left( \lambda^{(E_n)} \left( \int E_n(r) dr - 1 \right) \right) \\
& + \beta^{(v)} \sum_n \left( \bar{v}_n \log \frac{\bar{v}_n}{v_n} - \bar{v}_n + v_n \right) \\
& + \beta^{(I)} \iint \left( M^{(I)}(f, r) \log \frac{M^{(I)}(f, r)}{\bar{M}^{(I)}(f, r)} - M^{(I)}(f, r) + \bar{M}^{(I)}(f, r) \right) df dr \\
& + \beta^{(E)} \int \left( \bar{E}(r) \log \frac{\bar{E}(r)}{E_n(r)} - \bar{E}(r) + E_n(r) \right) dr
\end{aligned}$$

10

【0 1 3 9】

上記コスト関数は、上記 [ 数 6 ] に示したコスト関数とは、以下の 2 点で相違する。

20

【0 1 4 0】

1. コスト関数に倍音ピークの相対強度  $v_n$  と制約パラメータ  $v_n$  との独立性を示す距離が付加されている。

【0 1 4 1】

2. 時間方向エンベロープの制約パラメータ  $E(r)$  が平均の時間方向エンベロープとは異なったものとなっている。

【0 1 4 2】

制約パラメータ  $v_n$  はオン・オフセット区間のスペクトログラムに対してのみ上記コスト関数を最小化することによって得られるパラメータである。 $v_n$  は次式より得られる。

【数 3 0】

30

$$\bar{v}_n = \frac{\int_{r_{\text{on}}}^{r_{\text{off}}} \int S_n^{(H)}(f, r) df dr}{\sum_n \int_{r_{\text{on}}}^{r_{\text{off}}} \int S_n^{(H)}(f, r) df dr}$$

【0 1 4 3】

さらに、倍音ピークの相対強度に関する制約コストの付加より、倍音ピークの相対強度の更新式は次式のように改訂される。

【数 3 1】

40

$$v_n = \frac{\iint S_n^{(H)}(f, r) df dr + \beta^{(v)} \bar{v}_n}{\sum_n \iint S_n^{(H)}(f, r) df dr + \beta^{(v)}}$$

【0 1 4 4】

また、時間方向のエンベロープに関する制約パラメータ  $E(r)$  は次の式から求められる。

。

【数 3 2】

$$\bar{E}(r) = w(r) \sum n E_n(r) / N$$

【0 1 4 5】

これらの式を用いると、さらに精度の高い音色の変更（操作）が可能になる。

【数 3 3】

$$\begin{cases} w(r) = \exp \left[ -\frac{(r - r_{\text{on}})^2}{2\psi^2} \right] & (r < r_{\text{on}}) \\ w(r) = 1 & (r_{\text{on}} \leq r \leq r_{\text{off}}) \\ w(r) = \exp \left[ -\frac{(r - r_{\text{off}})^2}{2\psi^2} \right] & (r_{\text{off}} < r) \end{cases} \quad 10$$

【0 1 4 6】

なお音高軌跡の更新式は、下記のようになる。

【数 3 4】

$$\mu(r) = \frac{\sum_n n \int S_n^{(H)}(f, r) f \sqrt{1 + Bn^2} df}{\sum_n n^2 \int S_n^{(H)}(f, r) (1 + Bn^2) df} \quad 20$$

【0 1 4 7】

また非調和度の更新式は、下記のようになる。

【数 3 5】

$$B = \frac{2 \sum_n \iint S_n^{(H)}(f, r) n^3 \mu(r) (f - n\mu(r)) df dr}{\sum_n \iint S_n^{(H)}(f, r) (n^3 \mu(r))^2 df dr} \quad 30$$

【0 1 4 8】

さらに時間方向のエンベロープの更新式は下記のようになる。

【数 3 6】

$$E_n(r) = \frac{\int S_n^{(H)}(f, r) dr + \beta^{(E)} \bar{E}(r)}{\iint w^{(H)} F_n(f, r) df dr + \beta^{(E)}} \quad 40$$

【0 1 4 9】

上記実施の形態では、音高、音長、音色、楽譜の操作を行って第 1 の種類の楽器の単音を第 2 の種類の楽器の単音に置換して、しかも第 1 の種類の楽器で未知の楽譜を演奏した場合の音楽音響信号を生成することができる。しかしながら、本発明は第 1 の種類の楽器を用いて未知の楽譜を演奏した場合の音楽音響信号を生成する場合にも当然にして適用することができる。

【産業上の利用可能性】

【 0 1 5 0 】

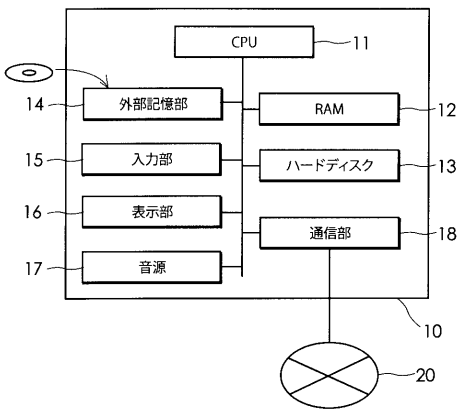
本発明によれば、調波モデルを構成するパラメータのうち音色に係わるパラメータの置換（変更）により、音色の変更（操作）をすることができるので、簡単に各種の音色変更を実現することができる。

【 符号の説明 】

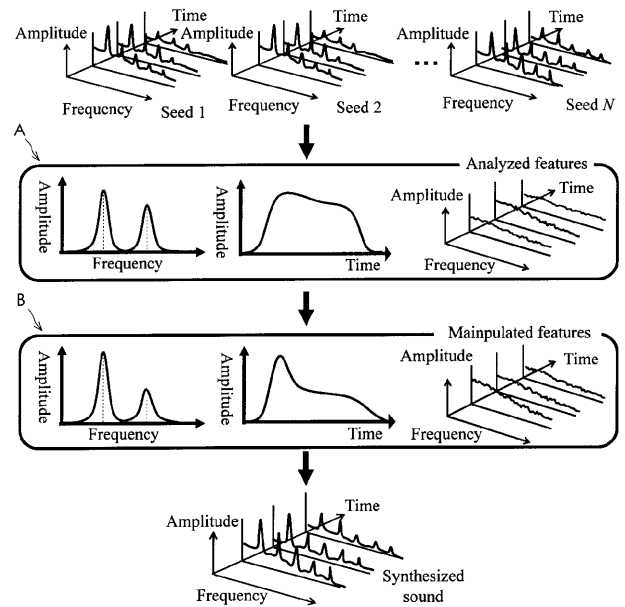
【 0 1 5 1 】

- 1 音響信号分離部
- 2 信号抽出保存部
- 3 分離音響信号分析保存部
- 4 置換パラメータ作成保存部
- 5 楽器分類判定部
- 6 置換用パラメータ保存部
- 7 合成分離音響信号生成部
- 8 信号加算部
- 9 A 音高操作部
- 9 B 音長操作部

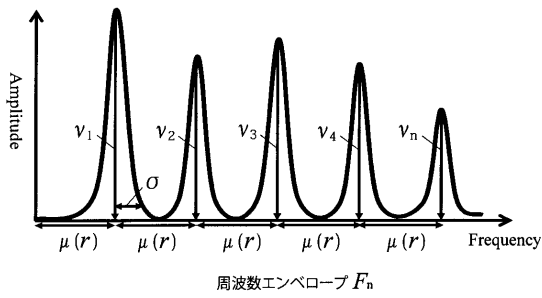
【 図 1 】



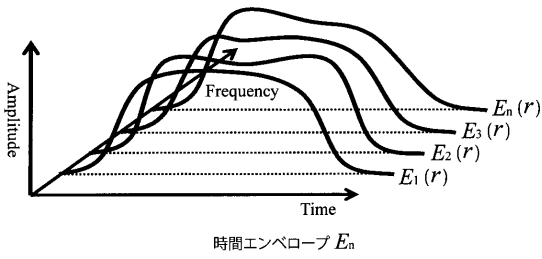
【 図 2 】



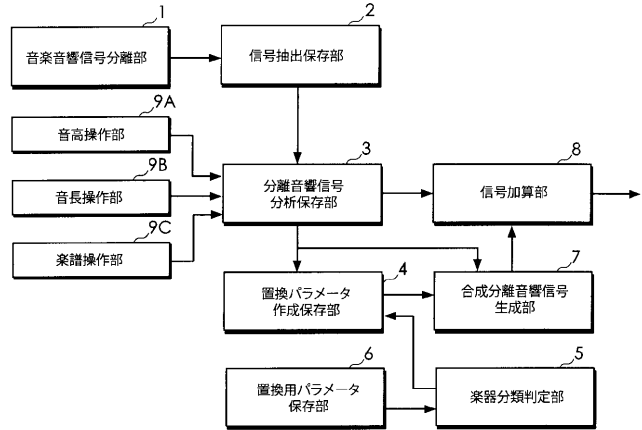
【 図 3 】



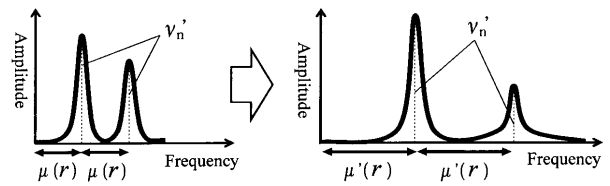
【 図 4 】



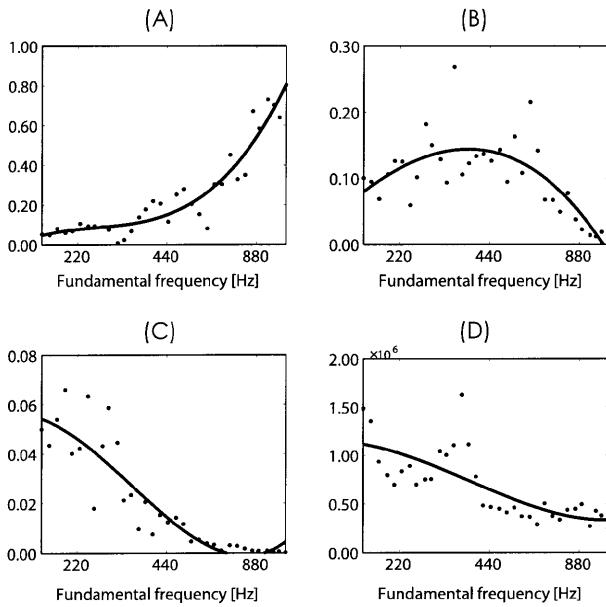
【 図 5 】



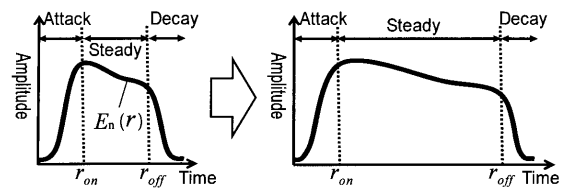
【 図 6 】



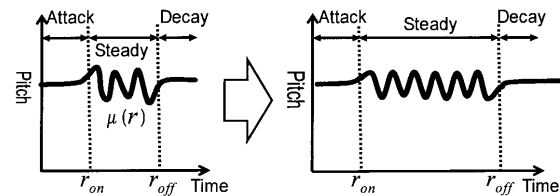
【 図 7 】



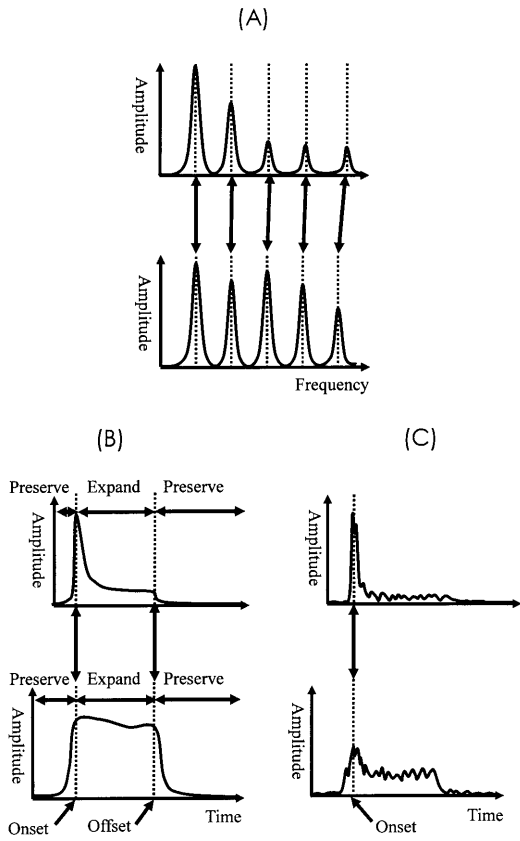
【 図 8 】



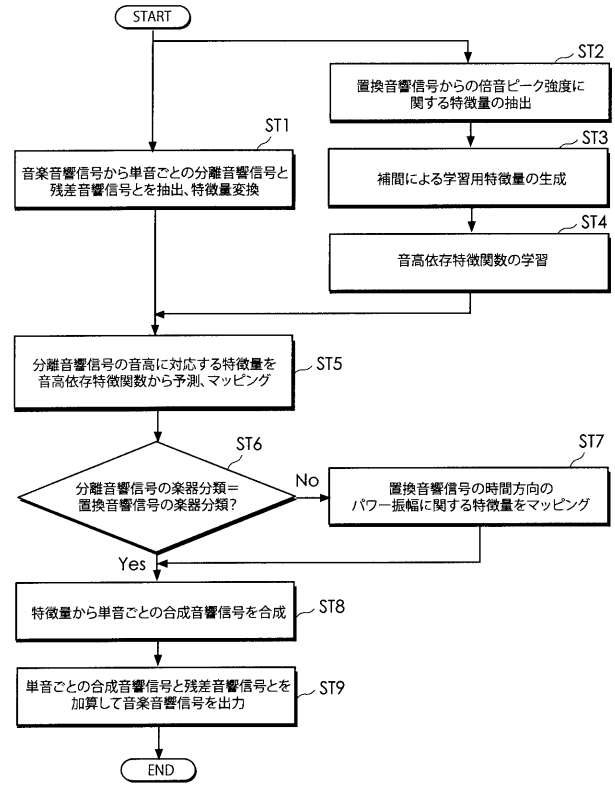
【 図 9 】



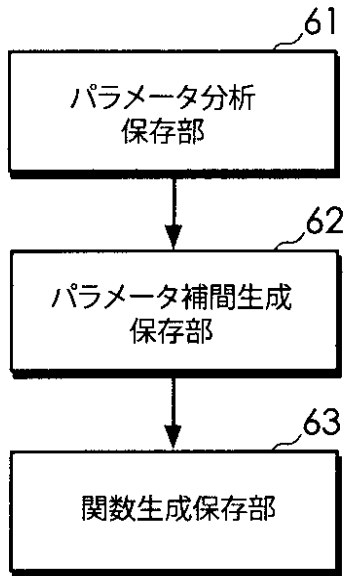
【図10】



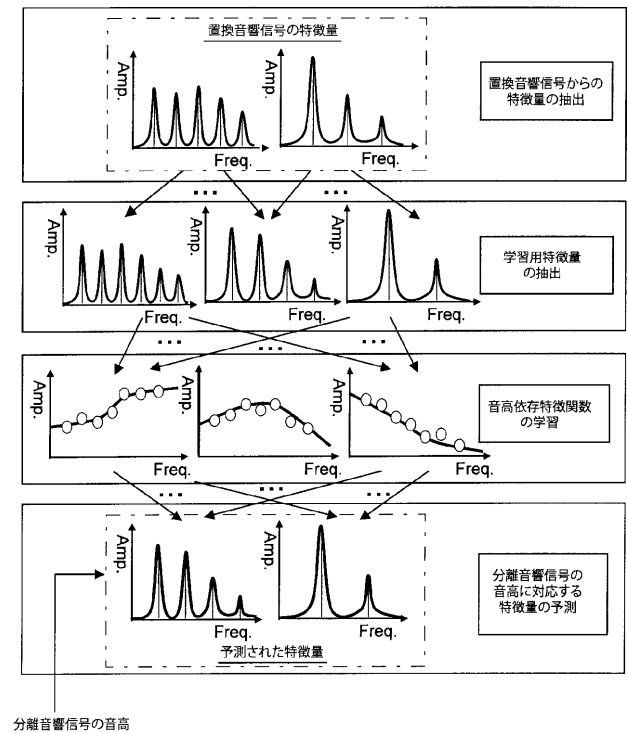
【図11】



【図12】

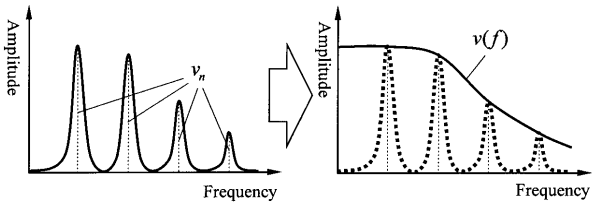


【図13】

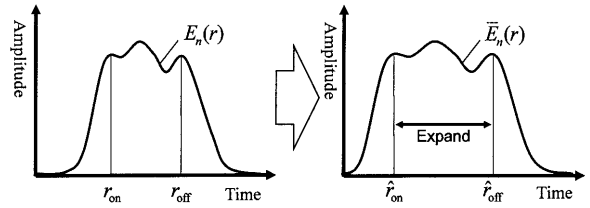




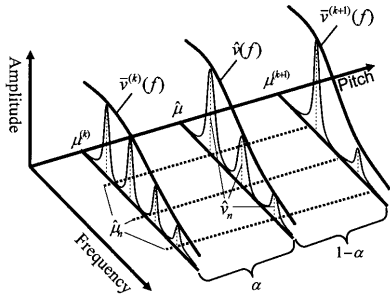
【 図 1 4 】



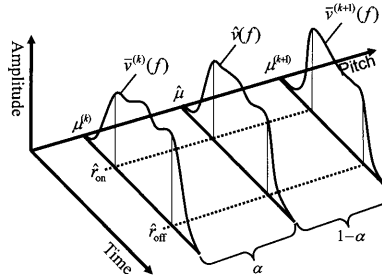
【 図 1 6 】



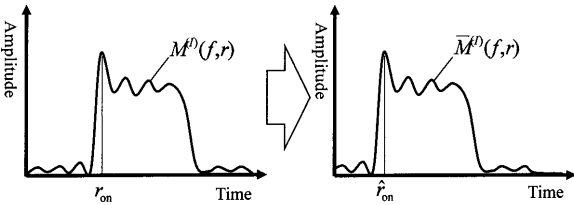
【 図 1 5 】



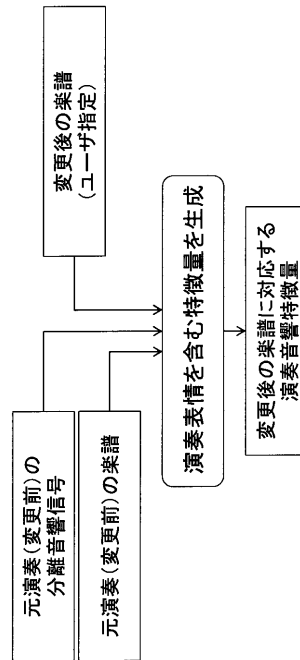
【 図 1 7 】



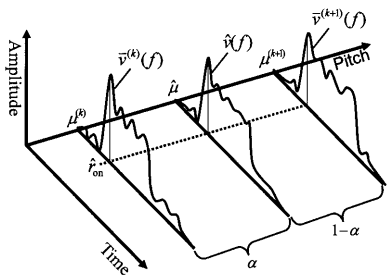
【 図 1 8 】



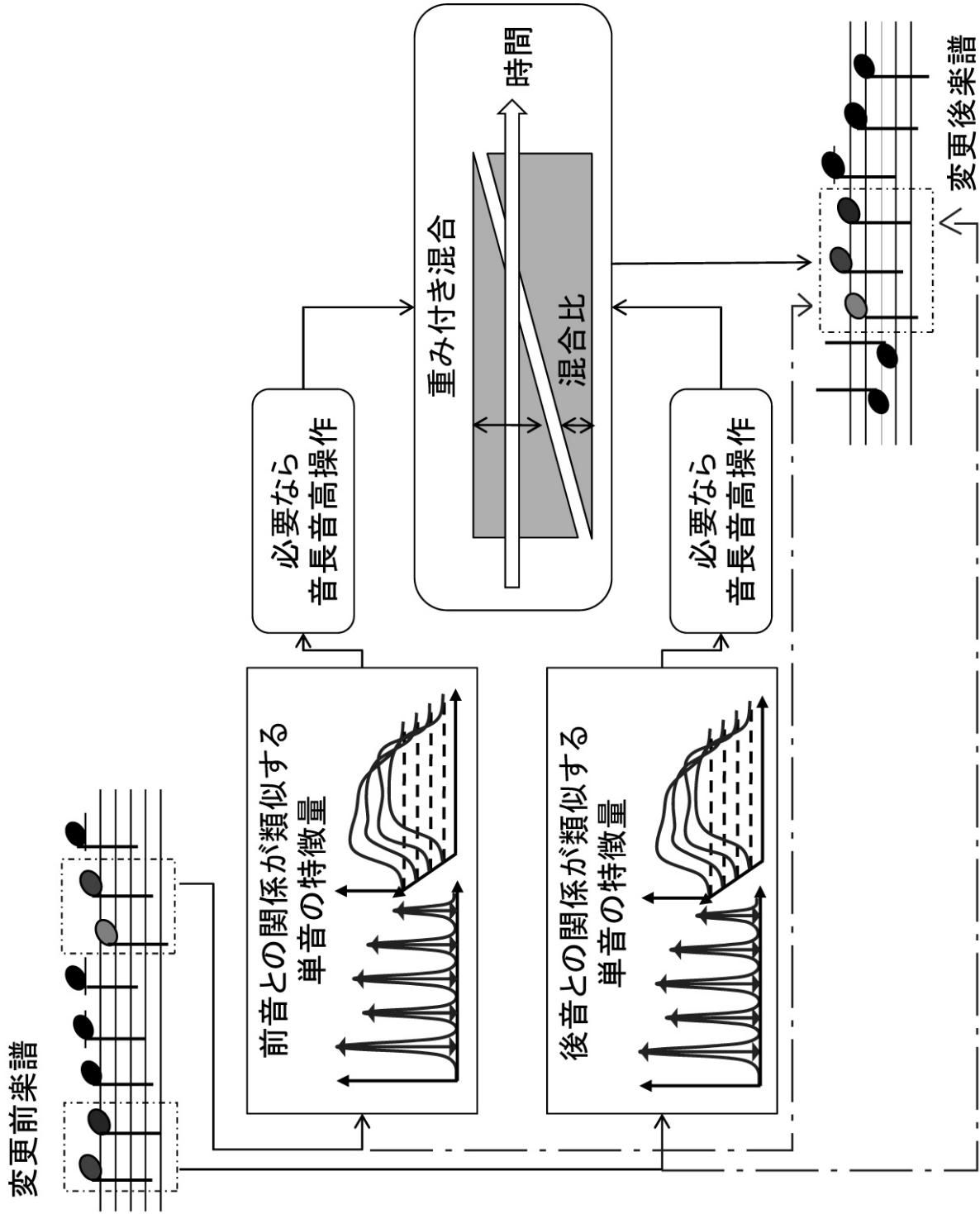
【 図 2 1 】



【 図 1 9 】



【 図 2 0 】



## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP2010/052293
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> G10L21/02(2006.01)i, G10H1/00(2006.01)i, G10H1/06(2006.01)i, G10L11/00(2006.01)i, G10L13/02(2006.01)i  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G10L21/02, G10H1/00, G10H1/06, G10L11/00, G10L13/02  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2010 Kokai Jitsuyo Shinan Koho 1971-2010 Toroku Jitsuyo Shinan Koho 1994-2010  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) JSTPlus (JDreamII), JST7580 (JDreamII)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 5-188931 A (Sony Corp.), 30 July 1993 (30.07.1993), entire text; all drawings (Family: none)	1-22
A	JP 2007-17818 A (Casio Computer Co., Ltd.), 25 January 2007 (25.01.2007), entire text; all drawings (Family: none)	1-22
A	JP 2006-5807 A (Kyoto University), 05 January 2006 (05.01.2006), entire text; all drawings & US 2005/0283361 A1	1-22
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 15 April, 2010 (15.04.10)		Date of mailing of the international search report 27 April, 2010 (27.04.10)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/052293

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2008/133097 A1 (Kyoto University), 06 November 2008 (06.11.2008), entire text; all drawings & EP 2148321 A	1-22
A	JP 2002-529773 A (Paul Reed Smith Guitars), 10 September 2002 (10.09.2002), entire text; all drawings & WO 2000/026897 A1 & EP 1125272 & DE 69904640 T2 & CN 1325525 A & KR 2001-082250	1-22

国際調査報告		国際出願番号 PCT/JP2010/052293									
A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. G10L21/02(2006.01)i, G10H1/00(2006.01)i, G10H1/06(2006.01)i, G10L11/00(2006.01)i, G10L13/02(2006.01)i											
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int.Cl. G10L21/02, G10H1/00, G10H1/06, G10L11/00, G10L13/02											
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2010年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2010年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2010年</td> </tr> </table>				日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2010年	日本国実用新案登録公報	1996-2010年	日本国登録実用新案公報	1994-2010年
日本国実用新案公報	1922-1996年										
日本国公開実用新案公報	1971-2010年										
日本国実用新案登録公報	1996-2010年										
日本国登録実用新案公報	1994-2010年										
国際調査で使用了電子データベース (データベースの名称、調査に使用した用語) JSTPlus(JDreamII), JST7580(JDreamII)											
C. 関連すると認められる文献											
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号									
A	JP 5-188931 A (ソニー株式会社) 1993.07.30, 全文、全図 (ファミリーなし)	1-22									
A	JP 2007-17818 A (カシオ計算機株式会社) 2007.01.25, 全文、全図 (ファミリーなし)	1-22									
A	JP 2006-5807 A (国立大学法人京都大学) 2006.01.05, 全文、全図 & US 2005/0283361 A1	1-22									
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。											
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願		の日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献									
国際調査を完了した日 15.04.2010		国際調査報告の発送日 27.04.2010									
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号		特許庁審査官 (権限のある職員) 鈴木 聡一郎	5Z 3864								
		電話番号 03-3581-1101	内線 3589								

国際調査報告		国際出願番号 PCT/JP2010/052293
C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2008/133097 A1 (国立大学法人京都大学) 2008.11.06, 全文、全図 & EP 2148321 A	1-22
A	JP 2002-529773 A (ポール リード スミス ギター) 2002.09.10, 全文、全図 & WO 2000/026897 A1 & EP 1125272 & DE 69904640 T2 & CN 1325525 A & KR 2001-082250	1-22

## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(72)発明者 奥乃 博

京都府京都市左京区吉田本町 国立大学法人京都大学大学院情報学研究科内

Fターム(参考) 5D378 AD52

(注)この公表は、国際事務局(WIPO)により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願(日本語実用新案登録出願)の国際公開の効果は、特許法第184条の10第1項(実用新案法第48条の13第2項)により生ずるものであり、本掲載とは関係ありません。