

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5654897号  
(P5654897)

(45) 発行日 平成27年1月14日(2015.1.14)

(24) 登録日 平成26年11月28日(2014.11.28)

(51) Int.Cl. F I  
**G 1 O G 3/04 (2006.01)** G 1 O G 3/04  
**G 1 O H 1/00 (2006.01)** G 1 O H 1/00 1 O 2 Z

請求項の数 6 (全 34 頁)

(21) 出願番号	特願2011-37526 (P2011-37526)	(73) 特許権者	000005326
(22) 出願日	平成23年2月23日(2011.2.23)		本田技研工業株式会社
(65) 公開番号	特開2011-180590 (P2011-180590A)		東京都港区南青山二丁目1番1号
(43) 公開日	平成23年9月15日(2011.9.15)	(74) 代理人	100064908
審査請求日	平成25年11月27日(2013.11.27)		弁理士 志賀 正武
(31) 優先権主張番号	61/309, 618	(74) 代理人	100108578
(32) 優先日	平成22年3月2日(2010.3.2)		弁理士 高橋 詔男
(33) 優先権主張国	米国 (US)	(74) 代理人	100146835
			弁理士 佐伯 義文
		(74) 代理人	100094400
			弁理士 鈴木 三義
		(74) 代理人	100107836
			弁理士 西 和哉
		(74) 代理人	100108453
			弁理士 村山 靖彦

最終頁に続く

(54) 【発明の名称】 楽譜位置推定装置、楽譜位置推定方法、及び楽譜位置推定プログラム

(57) 【特許請求の範囲】

【請求項1】

入力された音響信号の周波数領域の振幅である音響スペクトログラムを算出し、前記音響スペクトログラムから音階毎の強度を要素とする音響クロマベクトルを生成する音響特徴量生成部と、

単位時間毎の音階を表す楽譜情報から合成した音響信号の周波数特性の振幅である楽譜スペクトログラムを生成し、音階毎に前記楽譜情報が含む音階であるか否かを表す値を要素とする楽譜クロマベクトルを生成する楽譜特徴量生成部と、

仮想楽譜位置と仮想拍間隔の組毎の重み係数として、前記音響クロマベクトルと前記仮想楽譜位置における前記楽譜クロマベクトルとの相関性を示すクロマベクトル重み係数と、前記音響スペクトログラムと前記仮想楽譜位置における前記音響スペクトログラムとの相関性を示す周波数特性重み係数と、前記音響スペクトログラムと前記仮想拍間隔ずれた時刻の前記音響スペクトログラムとの相関性を示すテンポ重み係数との積を算出する重み算出部と、

前記組のうち前記重み算出部が算出した重み係数が所定の重み係数よりも大きい組を複製する再標本化部と、

前記組の仮想楽譜位置の分布から楽譜位置を推定する楽譜位置推定部と、  
を備えることを特徴とする楽譜位置推定装置。

【請求項2】

前記組毎に仮想テンポと第1の乱数に基づいて前記仮想楽譜位置を更新し、第2の乱数

に基づいて前記仮想テンポを更新する状態遷移部を備えることを特徴とする請求項 1 に記載の楽譜位置推定装置。

【請求項 3】

前記組毎の仮想楽譜位置に基づいて前記推定した楽譜位置の信頼度を算出し、前記信頼度に基づいて前記推定した楽譜位置を出力するか否かを判断する信頼度算出部を備えること、  
を特徴とする請求項 1 または請求項 2 に記載の楽譜位置推定装置。

【請求項 4】

少なくとも前記クロマベクトル重み係数と前記周波数特性重み係数のいずれかに基づいて楽譜特徴量と音響特徴量の類似性を示す信頼度を算出する信頼度算出部を備え、  
前記重み算出部は、  
前記算出された信頼度に基づき前記重み係数を算出する音響信号の区間を延長するか否かを判断すること、  
を特徴とする請求項 1 または請求項 2 に記載の楽譜位置推定装置。

【請求項 5】

楽譜位置推定装置における楽譜位置推定方法において、  
入力された音響信号の周波数領域の振幅である音響スペクトログラムを算出し、前記音響スペクトログラムから音階毎の強度を要素とする音響クロマベクトルを生成する音響特徴量生成過程と、  
単位時間毎の音階を表す楽譜情報から合成した音響信号の周波数特性の振幅である楽譜スペクトログラムを生成し、音階毎に前記楽譜情報が含む音階であるか否かを表す値を要素とする楽譜クロマベクトルを生成する楽譜特徴量生成過程と、

仮想楽譜位置と仮想拍間隔の組毎の重み係数として、前記音響クロマベクトルと前記仮想楽譜位置における前記楽譜クロマベクトルとの相関性を示すクロマベクトル重み係数と、前記音響スペクトログラムと前記仮想楽譜位置における前記音響スペクトログラムとの相関性を示す周波数特性重み係数と、前記音響スペクトログラムと前記仮想拍間隔ずれた時刻の前記音響スペクトログラムとの相関性を示すテンポ重み係数との積を算出する重み算出過程と、

前記組のうち前記重み算出部が算出した重み係数が所定の重み係数よりも大きい組を複製する再標準化過程と、  
前記組の仮想楽譜位置の分布から楽譜位置を推定する楽譜位置推定過程と、  
を有すること、  
を特徴とする楽譜位置推定方法。

【請求項 6】

楽譜位置推定装置のコンピュータに、  
入力された音響信号の周波数領域の振幅である音響スペクトログラムを算出し、前記音響スペクトログラムから音階毎の強度を要素とする音響クロマベクトルを生成する音響特徴量生成手順、  
単位時間毎の音階を表す楽譜情報から合成した音響信号の周波数特性の振幅である楽譜スペクトログラムを生成し、音階毎に前記楽譜情報が含む音階であるか否かを表す値を要素とする楽譜クロマベクトルを生成する楽譜特徴量生成手順、

仮想楽譜位置と仮想拍間隔の組毎の重み係数として、前記音響クロマベクトルと前記仮想楽譜位置における前記楽譜クロマベクトルとの相関性を示すクロマベクトル重み係数と、前記音響スペクトログラムと前記仮想楽譜位置における前記音響スペクトログラムとの相関性を示す周波数特性重み係数と、前記音響スペクトログラムと前記仮想拍間隔ずれた時刻の前記音響スペクトログラムとの相関性を示すテンポ重み係数との積を算出する重み算出手順、

前記組のうち前記重み算出部が算出した重み係数が所定の重み係数よりも大きい組を複製する再標準化手順、  
前記組の仮想楽譜位置の分布から楽譜位置を推定する楽譜位置推定手順、

を実行させるための楽譜位置推定プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、楽譜位置推定装置、楽譜位置推定方法、及び楽譜位置推定プログラムに関する。

【背景技術】

【0002】

近年、ロボットが人間社会においてますます利用されることが期待されている。このようなロボットは、人間とのインタラクション（相互作用；interaction）を行う手段を備えることが求められている。その手段を実現するためのメディア（媒体；medium）は、例えば言語による対話、音楽、アイコンタクト（視線を合わせること；eye contact）等である。とりわけ、音楽は、有望なメディアであり、人間社会において地域や人種を越えた重要かつ共通な要素である。例えば、ロボットが、音楽にあわせて人間と踊ったり、歌ったり、又は楽器を演奏することが期待されている。

【0003】

例えば、特許文献1には楽譜を2次元画像データに変換し、この画像データから五線データ及び小節線データを除去し、音符および記号を認識する画像認識装置が記載されている。また、特許文献2には、楽譜データおよびグルーピングされた構造分析結果に基づき、楽曲中の各拍点の局所的な強さの値を算出し、各拍点に対応づけてその強さの値を記憶手段に記憶し、次の階層の拍節構造のすべての候補において、その候補に含まれる拍点の強さの値を記憶手段から読み出して総和を求め、それぞれの候補における拍点の強さの値の総和を比較し、総和が最も大きい候補を次の階層の拍節構造として選択する楽曲の拍節構造の自動分析方法が提案されている。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特許第3147846号公報

【特許文献2】特開2006-201278号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかし、特許文献2記載の方法では、収録した楽曲が楽譜上のどの位置（以下、楽譜位置（Score Position）と呼ぶ）にあるかを認識できない。つまり、特許文献2記載の方法は、演奏されている曲のビート（拍；beat）時間やテンポの抽出に失敗することがあるという問題点があった。

【0006】

本発明は上記の点に鑑みてなされたものであり、演奏された音楽に対する楽譜位置を推定できる楽譜位置推定装置、楽譜位置推定方法、及び楽譜位置推定プログラムを提供することを課題としている。

【課題を解決するための手段】

【0007】

（1）本発明は上記の課題を解決するためになされたものであり、本発明の一態様は、入力された音響信号の周波数領域の振幅である音響スペクトログラムを算出し、前記音響スペクトログラムから音階毎の強度を要素とする音響クロマベクトルを生成する音響特徴量生成部と、単位時間毎の音階を表す楽譜情報から合成した音響信号の周波数特性の振幅である楽譜スペクトログラムを生成し、音階毎に前記楽譜情報が含む音階であるか否かを表す値を要素とする楽譜クロマベクトルを生成する楽譜特徴量生成部と、仮想楽譜位置と仮想拍間隔の組毎の重み係数として、前記音響クロマベクトルと前記仮想楽譜位置における前記楽譜クロマベクトルとの相関性を示すクロマベクトル重み係数と、前記音響スペクト

10

20

30

40

50

ログラムと前記仮想楽譜位置における前記音響スペクトログラムとの相関性を示す周波数特性重み係数と、前記音響スペクトログラムと前記仮想拍間隔ずれた時刻の前記音響スペクトログラムとの相関性を示すテンポ重み係数との積を算出する重み算出部と、前記組のうち前記重み算出部が算出した重み係数が所定の重み係数よりも大きい組を複製する再標本化部と、前記組の仮想楽譜位置の分布から楽譜位置を推定する楽譜位置推定部と、を備えることを特徴とする楽譜位置推定装置である。

【0008】

(2) 本発明の他の態様は、前記組毎に仮想テンポと第1の乱数に基づいて前記仮想楽譜位置を更新し、第2の乱数に基づいて前記仮想テンポを更新する状態遷移部を備えることを特徴とする(1)の楽譜位置推定装置である。

10

【0010】

(3) 本発明の他の態様は、前記組毎の仮想楽譜位置に基づいて前記推定した楽譜位置の信頼度を算出し、前記信頼度に基づいて前記推定した楽譜位置を出力するか否か判断する信頼度算出部を備えること、を特徴とする(1)又は(2)の楽譜位置推定装置である。

【0011】

(4) 本発明の他の態様は、少なくとも前記クロマベクトル重み係数と前記周波数特性重み係数のいずれかに基づいて楽譜特徴量と音響特徴量の類似性を示す信頼度を算出する信頼度算出部を備え、前記重み算出部は、前記算出された信頼度に基づき前記重み係数を算出する音響信号の区間を延長するか否か判断すること、を特徴とする(1)又は(2)の楽譜位置推定装置である。

20

【発明の効果】

【0012】

(1)に記載した態様によれば、入力された音響信号の特徴量を生成し、音響信号の特徴量と楽譜情報の特徴量に基づき重み係数を算出し、重み係数に対応する仮想楽譜位置及び仮想テンポを用いて楽譜位置を推定するので、演奏された音楽に対する楽譜位置を精度よく推定することが可能になる。また、クロマベクトル又は周波数特性を含む特徴量に基づき音響信号と楽譜情報との類似性を示す重み係数を算出するため、音響信号と楽譜情報が類似する仮想楽譜位置及び仮想テンポが重視される。そのため、当該発明は演奏された音楽に対する楽譜位置を精度よく推定することが可能になる。

(2)に記載した態様によれば、仮想テンポと第1の乱数に基づいて仮想楽譜位置を更新し、第2の乱数に基づいて仮想テンポを更新するため、演奏された音楽のテンポが変動しても楽譜位置を頑健に推定することが可能になる。

30

(3)に記載した態様によれば、推定した楽譜位置の信頼度に基づいて推定した楽譜位置を出力するか否か判断するため、出力された楽譜位置の信頼度を確保することが可能になる。

(4)に記載した態様によれば、算出された信頼度に基づき重み係数を算出する特徴量の区間を延長するか否か判断するため、推定した楽譜位置の信頼度が確保される特徴量を含むように区間を延長して、出力された楽譜位置の信頼度を確保することが可能になる。

【図面の簡単な説明】

【0013】

【図1】本発明の第1の実施形態に係る楽譜位置推定装置の構成を示す概略図である。

【図2】本実施形態に係る楽譜位置推定処理の概要を示す概念図である。

【図3】本実施形態に係る楽譜位置推定処理の概要を示すフローチャートである。

【図4】本実施形態に係る楽譜位置推定処理の一部を示すフローチャートである。

【図5】本実施形態に係る楽譜位置推定処理の他の一部を示すフローチャートである。

【図6】本実施形態に係る楽譜位置推定処理の残りの一部を示すフローチャートである。

【図7】本実施形態に係る楽譜位置の推定誤差の一例を示す図である。

【図8】本実施形態に係る信頼度の一例を示す図である。

【図9】本実施形態に係る推定誤差が生じる割合の一例を示す図である。

【図10】本実施形態に係る推定誤差の一例を示す図である。

40

50

【図 1 1】本実施形態に係る推定誤差が生じる割合の他の例を示す図である。  
 【図 1 2】本実施形態に係る推定誤差の他の例を示す図である。  
 【図 1 3】本実施形態に係る推定誤差と粒子数の関係の一例を示す図である。  
 【図 1 4】本実施形態に係る精度と再現率の一例を示す図である。  
 【図 1 5】本発明の第 2 の実施形態に係る楽譜位置推定装置の構成を示す概略図である。  
 【図 1 6】本実施形態に係る観測時間を変更する処理の概要を示す概念図である。  
 【図 1 7】本実施形態に係る観測時間を変更する処理を示すフローチャートである。  
 【図 1 8】本実施形態に係る楽譜位置推定装置が行う D T W 法を示すフローチャートである。

【図 1 9】本実施形態に係る楽譜特徴量と音響特徴量とのマッチング処理を示すフローチャートである。 10

【発明を実施するための形態】

【0014】

(第 1 の実施形態)

以下、図面を参照しながら本発明の第 1 の実施形態について説明する。

図 1 は、本実施形態に係る楽譜位置推定装置 1 の構成を示す概略図である。楽譜位置 (score position) 推定装置 1 は、音響信号入力部 101、音響特徴量生成部 102、楽譜情報記憶部 103、楽譜情報入力部 104、楽譜特徴量生成部 105、楽譜位置推定部 106、信頼度 (confidence) 算出部 107、楽譜位置出力部 108、及びテンポ出力部 109 を含んで構成される。 20

【0015】

音響信号入力部 101 は、演奏された音楽による音波をデジタル音響信号に変換し、変換したデジタル音響信号を音響特徴量生成部 102 に出力する。音響信号入力部 101 は、例えばマイクロホンとアナログ・デジタル (Analog - to - Digital ; A / D) 変換部 (図示せず) とを含んで構成される。このマイクロホンは、人間が聴取することができる周波数帯域 (例えば、20 Hz - 20 kHz) の音波を受信してアナログ音響信号を生成し、生成したアナログ音響信号を A / D 変換部に出力する。A / D 変換部は、マイクロホンから入力されたアナログ音響信号をデジタル音響信号に変換し、変換したデジタル音響信号を音響特徴量生成部 102 に出力する。ここで、A / D 変換部は、入力されたアナログ音響信号を、例えば、サンプリング周波数 44.1 kHz、振幅を 16 ビットの 2 進 (バイナリ ; binary) データにパルス符号変調 (Pulse Code Modulation ; PCM) することで、量子化されたデジタル音響信号に変換する。 30

【0016】

音響特徴量生成部 102 は、音響信号入力部 101 から入力されたデジタル音響信号から、その特徴量 (音響特徴量) を生成し、生成した音響特徴量を楽譜位置推定部 106 に出力する。

具体的には、音響特徴量生成部 102 は、音響特徴量として、例えば、スペクトログラム (spectrogram) 及びクロマベクトル (chroma vector) を生成する。音響特徴量生成部 102 は、周波数特性分析部 1021 及びクロマベクトル生成部 1022 を含んで構成される。 40

周波数特性分析部 1021 は、音響信号入力部 101 から入力されたデジタル音響信号を時間領域信号から周波数領域信号に変換する。ここで、周波数特性分析部 1021 は、例えば、2048 個のサンプルからなるフレームごとにデジタル音響信号を高速フーリエ変換 (Fast Fourier Transform ; FFT) して周波数領域信号に変換する。このフレームは、FFT を行うデジタル音響信号の範囲である。周波数特性分析部 1021 は、予め設定された時刻 t 毎 (例えば、10 ms 間隔) に新たな信号のサンプル (例えば、サンプリング周波数が 44.1 kHz の場合、441 サンプル) が含まれるようにフレームをシフトさせる。

周波数特性分析部 1021 は、フレーム時刻 t 毎に変換した周波数領域信号の振幅 (絶 50

対値)をスペクトログラム(以下、音響スペクトログラムと呼ぶ) $X_{f,t}$ として算出し、算出した音響スペクトログラム $X_{f,t}$ をクロマベクトル生成部1022及び楽譜位置推定部106に出力する。なお、 $f$ は周波数を表す。

【0017】

クロマベクトル生成部1022は、周波数特性分析部1021から入力された音響スペクトログラム $X_{f,t}$ に基づいて音響信号によるクロマベクトル(以下、音響クロマベクトルと呼ぶ) $c_t^a$ を算出する。クロマベクトルとは、音階(chroma)毎の成分のパワーを要素として含むベクトルである。クロマベクトルは、例えば、西洋音楽を構成する12音階(C、C#、D、D#、E、F、F#、G、G#、A、A#、B)における音階毎の成分の強度を要素とするベクトル(例えば、要素数が12)である。

10

クロマベクトル生成部1022は、音響クロマベクトル $c_t^a = [c_t^a(1), c_t^a(2), \dots, c_t^a(12)]$ の音階 $n$ 毎の要素 $c_t^a(n)$ を、例えば、次式を用いて算出する。

【0018】

【数1】

$$c_t^a(n) = \sum_{h=Oct_L}^{Oct_H} \int_0^{\infty} BPF_{n,h}(f) |X_{f,t}|^2 df \quad (1)$$

20

【0019】

式(1)において、 $Oct_L$ は予め設定されたオクターブの下限(下限オクターブ)を表す。 $Oct_H$ は、予め設定されたオクターブの上限(上限オクターブ)を表す。 $BPF_{n,h}(f)$ は、 $h$ 番目のオクターブにおける音階 $n$ の成分を抽出する帯域通過フィルタ(バンドパスフィルタ、Band Pass Filter; BPF)の入出力特性を表す。 $h$ 番目のオクターブにおける音階 $n$ の周波数 $f_{n,h}$ において、その値は最大となり、周波数 $f$ がゼロ又は無限大に近づくと、ゼロに近似する関数である。また、 $|X_{f,t}|^2$ は、パワースペクトルを示す。楽譜位置推定装置1は、このクロマベクトルを用いることで、オクターブが異なっても音階毎の強度を抽出でき、音響信号の調波構造(harmonic structure)を無視することができる。よって、楽譜位置推定装置1は、旋律を表すために重要な音階の計時変化を抽出でき、旋律を表すために余り重要ではないオクターブの差異を除外することができる。

30

クロマベクトル生成部1022は、算出した音響クロマベクトル $c_t^a$ を楽譜位置推定部106に出力する。

【0020】

楽譜情報記憶部103は、楽譜情報を楽曲毎に記憶する。楽譜情報は、楽譜フレーム(score frame) $k$ ( $k$ は1又は1より大きい整数)毎に、楽曲を構成する音階を示す音階情報 $n_k^r$ を要素とするベクトル $n_k = [n_k^1, n_k^2, \dots, n_k^{r_k}]^T$ で表される。ここで、 $r$ は1又は1より大きく $r_k$ 又は $r_k$ よりも小さい整数である。 $r_k$ は、楽譜フレーム $k$ における音階の数を表す。 $T$ は、ベクトルの転置を表す。楽譜フレームとは、1つの楽曲の楽譜情報において楽曲を時間分割する単位時間である。例えば、四分音符の長さが12フレームである場合、楽譜情報の時間分解能(resolution)は十六分音符の三分の一となる。例えば、楽譜フレーム2において、音階C、E及びGを表す楽譜情報 $n_2$ は、要素数が3個のベクトル $[C, E, G]$ となる。

40

【0021】

楽譜情報入力部104は、楽譜情報記憶部103から、処理の対象とする楽曲の楽譜情報 $n_k$ を読み出す。楽譜情報入力部104は、読み出した楽譜情報 $n_k$ を楽譜特徴量生成部105に出力する。

【0022】

楽譜特徴量生成部105は、楽譜情報入力部104から入力された楽譜情報 $n_k$ に基づ

50

いて、その特徴量（以下、楽譜特徴量と呼ぶ）を生成し、生成した楽譜特徴量を楽譜位置推定部106に出力する。楽譜特徴量生成部105は、楽譜特徴量として、例えば、スペクトログラム及びクロマベクトルを生成する。ここで、楽譜特徴量生成部105が生成したスペクトログラム及びクロマベクトルを、それぞれ楽譜スペクトログラム及び楽譜クロマベクトルと呼ぶ。

楽譜特徴量生成部105は、周波数特性合成部1051及びクロマベクトル生成部1052を含んで構成される。

周波数特性合成部1051は、楽譜情報入力部から入力された楽譜情報 $n_k$ に基づき、楽譜スペクトログラム $X'_{f,k}$ を生成する。楽譜スペクトログラム $X'_{f,k}$ とは、楽譜情報 $n_k$ に基づき演奏されるべき音楽を示す音響信号の周波数特性である。

周波数特性合成部1051は、例えば次式のように調和混合正規分布モデル(harmonic gaussian mixture model; harmonic GMM)に基づいて楽譜スペクトログラム $X'_{f,k}$ を生成する。

【0023】

【数2】

$$X'_{f,k} = \sum_{r=1}^{n_k} \sum_{g=1}^G h(g) N(f; gF_{n_k}, \sigma^2) + C(f) \quad (2)$$

【0024】

式(2)の右辺第1項において、 $g$ は、調波インデックス(harmonic index)即ち、基本周波数(fundamental frequency) $F$ の整数倍(integer multiple)の周波数の成分を示すインデックスである。 $g=1$ は、基本波(fundamental tone)を示す。 $G$ は、調波の数(例えば10)を示す。 $h(g)$ は、第 $g$ 高調波( $g$ -th harmonic)の振幅(例えば、0.2の $g$ 乗)を表す。 $N(f; gF_{n_k}, \sigma^2)$ は、パワーの周波数特性が、中心周波数 $gF_{n_k}$ 、分散 $\sigma^2$ の正規分布(normal distribution, gaussian distribution)関数を表す。即ち、式(2)の右辺第1項は純粋な調和混合正規分布モデルに基づき、各音階を構成する調波成分(harmonic)を線形結合した成分を示す。なお、分散 $\sigma^2$ は、例えば0.8である。

【0025】

式(2)の右辺第2項において、 $C(f)$ は、ピンクノイズ(pink noise)成分のパワー(power)の周波数特性を示す。ピンクノイズとは、オクターブ帯域毎のパワーが等しい周波数特性を有しランダムな波形を表す信号系列である。式(2)の右辺において、 $C(f)$ を加算するのは、右辺第1項の成分が表すスペクトログラムを処理すると生じるゼロ除算(zero divide)を避けるためである。

$C(f)$ は、例えば次式(3)で示される周波数特性を示す。

【0026】

【数3】

$$C(f) = A \exp(-\alpha f) \quad (3)$$

【0027】

$A$ は、式(2)の右辺第2項の成分について右辺第1項の成分とのパワー比(例えば5%)を示す定数である。 $\alpha$ は、パワーの減少率を示す定数である。 $\alpha$ は、例えば、 $\log_{10}(C(f+\Delta f)/C(f)) = -0.6$ となる値である。但し、 $\Delta f$ は、帯域幅1kHzに対応する周波数領域信号を構成する周波数変数(frequency bin)の数である。

周波数特性合成部 1051 は、生成した楽譜スペクトログラム  $X'_{f,k}$  を楽譜位置推定部 106 に出力する。

【0028】

クロマベクトル生成部 1052 は、楽譜情報入力部 104 から入力された楽譜情報  $n_k$  に基づき楽譜クロマベクトル  $c_k^s$  を生成する。クロマベクトル生成部 1052 は、例えば、楽譜情報  $n_k$  に含まれる音階  $n$  に対応する要素  $c_k^s(n)$  を 1 とし、楽譜情報  $n_k$  に含まれない音階  $n$  に対応する要素  $c_k^s(n)$  をゼロとした楽譜クロマベクトルを生成する。例えば、 $n_2$  が [C, E, G] であるとき、クロマベクトル生成部 1052 は、楽譜クロマベクトル  $c_2^s = [1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0]$  を生成する。

10

クロマベクトル生成部 1052 は、生成した楽譜クロマベクトル  $c_k^s$  を楽譜位置推定部 106 に出力する。

【0029】

楽譜位置推定部 106 は、重み (weight) 算出部 1061、再標本化 (resampling) 部 1062、状態遷移 (state transition) 部 1063、楽譜位置算出部 1064 及びテンポ算出部 1065 を含んで構成される。

重み算出部 1061 は、音響特徴量生成部 102 から入力された音響特徴量と楽譜特徴量に基づいて後述する粒子毎に重み係数を算出する。重み算出部 1061 は、算出した粒子毎の重み係数を再標本化部 1062 に出力する。

【0030】

20

具体的には、重み算出部 1061 は、クロマベクトル重み (chroma vector weight) 算出部 10611、周波数特性重み (spectrogram weight) 算出部 10612、テンポ重み (tempo weight) 算出部 10613、及び粒子重み算出部 10614 を含んで構成される。

【0031】

クロマベクトル重み算出部 10611 は、クロマベクトル生成部 1022 から入力された音響クロマベクトル  $c_t^a$  とクロマベクトル生成部 1052 から入力された楽譜クロマベクトル  $c_k^s$  に基づいてクロマベクトル重み係数  $w^{ch}$  を算出する。クロマベクトル重み係数  $w^{ch}$  は、音響クロマベクトル  $c_t^a$  と楽譜クロマベクトル  $c_k^s$  との類似度又は相関性を示す尺度である。クロマベクトル重み算出部 10611 は、例えば、以下の式を用いてクロマベクトル重み係数  $w^{ch}$  を算出する。

30

【0032】

【数 4】

$$w^{ch} = \sum_{\tau=t-L}^t c_{\tau}^a \cdot c_{k_{\tau}^i}^s / L \quad (4)$$

【0033】

式 (4) において、 $\tau$  は時刻を表す。L は予め設定された時間を表す。t は最新のフレーム時刻を表す。k は、観測時刻における楽譜フレーム k である。k と観測時刻は次式 (5) の関係がある。

40

【0034】

【数 5】

$$k_{\tau} = \text{floor}(k' - (k - \tau) / b' + 0.5) \quad (5)$$

【0035】

式 (5) において、 $\text{floor}(\dots)$  は、 $\text{floor}$  関数 (floor func

50



tion ; 床関数)を表す。即ち、 $floor(\dots)$ は、実数 $\dots$ を、その整数部に切り捨てる関数である。 $k'$ は、仮想楽譜位置である。 $b'$ は、仮想拍間隔(beat interval)である。 $k'$ 及び $b'$ については、後述する。

即ち、式(4)は、クロマベクトル重み係数 $w^{ch}$ が、音響クロマベクトル $c_t^a$ と楽譜クロマベクトル $c_{k_t}^s$ との内積を最新のフレーム $t$ から予め設定された時間 $L$ だけ遡って累積した値であることを表す。以下の説明では、この時間 $L$ を観測時間(窓長; window length)と呼ぶ。なお、 $L$ は、例えば2.5秒である。

なお、式(4)によれば、クロマベクトル重み係数 $w^{ch}$ は、音響クロマベクトル $c_t^a$ と楽譜クロマベクトル $c_{k_t}^s$ が類似する(つまり、両ベクトルの方向が近くなる)ほど増加し、音響クロマベクトル $c_t^a$ と楽譜クロマベクトル $c_{k_t}^s$ が一致するとき最大値1となる。

クロマベクトル重み算出部10611は、算出したクロマベクトル重み係数 $w^{ch}$ を粒子重み算出部10614に出力する。

【0036】

なお、クロマベクトル重み算出部10611は、後述する粒子(particle) $i$ 毎に定められた仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ が、状態遷移部1063から入力される。クロマベクトル重み算出部10611は、観測時間の始点 $t-L$ と終点 $t$ を、各々先行時間(look ahead time) $T$ だけ後の時刻に変更する。先行時間 $T$ とは、状態遷移部1063において1回の状態遷移毎に観測時刻を進める時間であり、例えば1秒である。これにより、クロマベクトル重み算出部10611は、観測時間長 $L$ を変更せずに観測時間をシフトさせる。また、クロマベクトル重み算出部10611は、上述の処理において粒子 $i$ 毎の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ に基づきクロマベクトル重み係数 $w_i^{ch}$ を算出及び出力する。

【0037】

周波数特性重み算出部10612は、周波数特性分析部1021から入力された音響スペクトログラム $X_{f,t}$ と周波数特性合成部1051から入力された楽譜スペクトログラム $X'_{f,k}$ に基づいて周波数特性重み係数 $w^{sp}$ を算出する。周波数特性重み係数 $w^{sp}$ は、音響スペクトログラム $X_{f,t}$ と楽譜スペクトログラム $X'_{f,k}$ との類似度又は相関性を示す尺度である。周波数特性重み算出部10612は、例えば、音響スペクトログラム $X_{f,t}$ と楽譜スペクトログラム $X'_{f,k}$ のカルバック・ライブラー情報量(Kullback-Leibler Divergence)に基づいて周波数特性重み係数 $w^{sp}$ と定める。ここで、周波数特性重み算出部10612は、次式(6)を用いて周波数特性重み係数 $w^{sp}$ を算出する。

【0038】

【数6】

$$w^{sp} = (1+Q)\exp(-Q) \quad (6)$$

【0039】

式(6)において、 $Q$ は、以下の式で表される実数である。

【0040】

【数7】

$$Q = \frac{1}{L} \sum_{\tau=t-L}^t \sum_f X_{f,\tau} \log \frac{X_{f,\tau}}{X'_{f,k_\tau}} \quad (7)$$

【0041】

但し、(7)式において、周波数特性重み算出部10612は、次式(8)に示すように規格化した音響スペクトログラム $X_{f,t}$ と楽譜スペクトログラム $X'_{f,k_t}$ を用いる。

【0042】

【数8】

$$\sum_f X_{f,t} = \sum_f X'_{f,k_t} = 1 \quad (8)$$

【0043】

10

また、周波数特性重み算出部10612においても、式(5)に示される時刻と楽譜フレーム $k_t$ との関係を用いる。

式(6)及び(7)によれば、周波数特性重み係数 $w^{s,p}$ は、音響スペクトログラム $X_{f,t}$ と楽譜スペクトログラム $X'_{f,k_t}$ が類似するほど増加し、両者が一致するとき、最大値1となる。このように、周波数特性重み算出部10612は、周波数特性重み係数 $w^{s,p}$ を用いることで、クロマベクトルだけでは表現されない、入力音響信号や楽譜情報による周波数特性の広がりを考慮して確実に入力音響信号と楽譜情報との対応関係を把握することができる。

周波数特性重み算出部10612は、算出した周波数特性重み係数 $w^{s,p}$ を粒子重み算出部10614に出力する。

20

【0044】

なお、周波数特性重み算出部10612は、粒子 $i$ 毎に定められた仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ が、状態遷移部1063から入力される。このとき、周波数特性重み算出部10612は、観測時間の始点 $t-L$ と終点 $t$ を、各々先行時間 $T$ だけ後の時刻に変更することで、観測時間を後にシフトさせる。また、周波数特性重み算出部10612は、上述の処理において粒子 $i$ 毎の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ に基づき周波数特性重み係数 $w_i^{s,p}$ を算出及び出力する。

【0045】

テンポ重み算出部10613は、周波数特性分析部1021から入力された音響スペクトログラム $X_{f,t}$ からテンポ重み係数 $w^t$ を算出する。テンポ重み係数 $w^t$ は、入力音響信号が示す仮想拍間隔 $b'$ の確からしさを示す尺度である。

30

テンポ重み算出部10613は、例えば、音響スペクトログラム $X_{f,t}$ と仮想拍間隔 $b'$ だけずれた音響スペクトログラム $X_{f,t}$ の正規化相互相関(normalized cross correlation)をテンポ重み係数 $w^t$ として算出する。即ち、テンポ重み算出部10613は、次式(9)を用いてテンポ重み係数 $w^t$ を算出する。

【0046】

【数9】

$$w^t = \frac{\sum_{\tau=t-L}^t \sum_f X_{f,\tau} X_{f,\tau - \text{floor}(b'+0.5)}}{\sqrt{\sum_{\tau=t-L}^t \sum_f X_{f,\tau}^2 \sum_{\tau=t-L}^t \sum_f X_{f,\tau - \text{floor}(b'+0.5)}^2}} \quad (9)$$

40

【0047】

式(9)において、仮想拍間隔 $b'$ が先行時間 $T$ と等しいか、より大きいとき、テンポ重み算出部10613は、テンポ重み係数 $w^t$ をゼロと定める。

式(9)によれば、テンポ重み係数 $w^t$ は、現在の音響スペクトログラムと過去の音響スペクトログラムの相関性を表す。例えば、同じパワーの大きさで同じ音階が拍間隔 $b'$

50

ごとに繰り返し演奏されている入力音響信号については、テンポ重み係数  $w^t$  は、1となる。

テンポ重み算出部 10613 は、算出したテンポ重み係数  $w^t$  を粒子重み算出部 10614 に出力する。

【0048】

なお、テンポ重み算出部 10613 は、後述する粒子毎に定められた仮想拍間隔  $b_i'$  が、状態遷移部 1063 から入力される。このとき、テンポ重み算出部 10613 は、観測時間の始点  $t - L$  と終点  $t$  を、各々先行時間  $T$  だけ後の時刻に変更することで、観測時間がシフトさせる。また、テンポ重み算出部 10613 は、上述の処理において粒子  $i$  毎の仮想拍間隔  $b_i'$  に基づきテンポ重み係数  $w_i^t$  を算出及び出力する。

10

【0049】

粒子重み算出部 10614 は、クロマベクトル重み算出部 10611 から入力されたクロマベクトル重み係数  $w_i^{ch}$ 、周波数特性重み算出部 10612 から入力された周波数特性重み係数  $w_i^{sp}$ 、及びテンポ重み算出部 10613 から入力されたテンポ重み係数  $w_i^t$  に基づき粒子  $i$  毎の重み係数  $w_i$  を算出する。

粒子  $i$  とは、本実施形態において仮想的な変数である仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  及び観測値である重み係数  $w_i$ 、の組である。これらの粒子は、粒子フィルタリング (particle filtering) 法により観測値から仮想的な変数を推定するために用いられる。例えば、本実施形態は、粒子  $i$  毎の仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  に基づき、楽譜位置  $k'$ 、拍間隔  $b'$  の推定値を定める。

20

具体的には、粒子重み算出部 10614 は、粒子  $i$  毎の粒子重み係数  $w_i$  を次式 (10) に基づいて算出する。

【0050】

【数10】

$$w_i = w_i^{ch} \cdot w_i^{sp} \cdot w_i^t \quad (10)$$

【0051】

粒子重み算出部 10614 は、算出した粒子  $i$  毎の仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  及び粒子重み係数  $w_i$  を再標本化部 1062 に出力する。

30

【0052】

再標本化部 1062 は、粒子重み算出部 10614 から粒子  $i$  毎の仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  及び粒子重み係数  $w_i$  が入力される。再標本化部 1062 は、粒子  $i$  毎の粒子重み係数  $w_i$  に基づき、各粒子を再標本化する。再標本化部 1062 は、再標本化した粒子  $i$  毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  を状態遷移部 1063 に出力する。再標本化部 1062 は、再標本化に係る処理において粒子数を一定に保つために、棄却する粒子数と、分割又は複製によって新たに生成する粒子数を等しくする。

【0053】

再標本化部 1062 は、例えば SIR (sampling importance resampling) 法を用いて粒子を再標本化する。

40

ここで、再標本化部 1062 は、次の処理を行って粒子重み  $w_i$  が小さい粒子  $i$  を棄却する。このような粒子  $i$  に対応する、仮想楽譜位置  $k_i'$  や仮想拍間隔  $b_i'$  の信頼性が低いからである。

例えば、再標本化部 1062 は、粒子  $i$  毎の粒子重み係数を、次式 (11) を用いて規格化し、規格化粒子重み係数  $p_i$  を算出する。

【0054】

【数 1 1】

$$p_i = \frac{w_i}{\sum_{i=1}^N w_i} \quad (11)$$

【0055】

式(11)において、Nは粒子数(例えば、300個)を示す。

再標本化部1062は、規格化粒子重み係数 $p_i$ が予め設定された規格化粒子重み係数の閾値 $p_{th}$ よりも小さいか否か判断する。再標本化部1062は、規格化粒子重み係数 $p_i$ が閾値 $p_{th}$ よりも小さいと判断した場合、対応する粒子 $i$ を棄却する。即ち、再標本化部1062は、粒子 $i$ の規格化粒子重み係数 $p_i$ 、仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ を削除する。再標本化部1062は、規格化粒子重み係数 $p_i$ が閾値 $p_{th}$ よりも大きい又は等しいと判断した場合、対応する粒子 $i$ を維持する。また、再標本化部1062は、棄却された粒子数をカウントする。

10

【0056】

再標本化部1062は、次の処理を行って粒子重み $w_i$ が大きい粒子 $i$ を分割又は複製する。このような粒子 $i$ に対応する、仮想楽譜位置 $k_i'$ や仮想拍間隔 $b_i'$ の信頼性が高いからである。

例えば、再標本化部1062は、粒子 $i$ 毎の粒子重み係数を、次式(12)を用いて再規格化し、再規格化粒子重み係数 $q_i$ を算出する。

20

【0057】

【数 1 2】

$$q_i = N \cdot p_i \quad (12)$$

【0058】

再標本化部1062は、再規格化粒子重み係数 $q_i$ が1よりも大きい粒子について整数部分の値 $n_{qi}$ (以下、再規格化整数と呼ぶ)と小数部分の値 $s_{qi}$ (以下、再規格化小数と呼ぶ)を算出する。

30

再標本化部1062は、再規格化整数 $n_{qi}$ が2又は2より多い粒子 $i$ を分割して、新たに $n_{qi} - 1$ 個の粒子 $j$ ( $j = i$ )を各々生成する。即ち、再標本化部1062は、新たな粒子 $j$ に対応する仮想楽譜位置 $k_j'$ 及び仮想拍間隔 $b_j'$ を分割直前の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ と定める。再標本化部1062は、棄却された粒子数から新たに生成した粒子数 $n_{qi} - 1$ を順次差し引くことにより、所定の粒子数Nに不足する粒子数 $N_s$ (以下、不足粒子数と呼ぶ)を算出する。

不足粒子数がゼロに達した場合には、再標本化部1062は、後述する粒子を複製する処理を行わない。他方、不足粒子数が1以上である場合に、再標本化部1062は、粒子を複製する処理を行う。

そして、再標本化部1062は、新たに複製される粒子の総数が不足粒子数 $N_s$ となるように、再規格化小数 $s_{qi}$ に比例する確率(以下、複製確率と呼ぶ)で、複製対象の粒子 $i$ を決定する。例えば、再標本化部1062は、次式(13)で算出した複製確率 $s'_{qi}$ が、乱数 $r_i$ (但し、 $0 < r_i < 1$ )よりも小さい粒子を複製対象の粒子と定める。

40

【0059】

【数 1 3】

$$s'_{qi} = \frac{N_s \cdot s_{qi}}{\sum s_{qi}} \quad (13)$$

50

## 【 0 0 6 0 】

再標本化部 1 0 6 2 は、定めた粒子各々について、新たに 1 個の粒子 ( j i ) を複製する。即ち、再標本化部 1 0 6 2 は、新たな粒子 j に対応する仮想楽譜位置  $k_j'$  及び仮想拍間隔  $b_j'$  を複製直前の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  と定める。

再標本化部 1 0 6 2 は、上述の処理を完了して決定した粒子 i 毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  を状態遷移部 1 0 6 3 に出力する。

## 【 0 0 6 1 】

状態遷移部 1 0 6 3 は、再標本化部 1 0 6 2 から粒子 i 毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  を入力する。状態遷移部 1 0 6 3 は、仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  を、例えば式 ( 1 4 ) 及び式 ( 1 5 ) を用いて更新する。

10

## 【 0 0 6 2 】

## 【 数 1 4 】

$$k'_i \leftarrow k'_i + \Delta T / b'_i + u \quad (14)$$

## 【 0 0 6 3 】

## 【 数 1 5 】

$$b'_i \leftarrow b'_i + v \quad (15)$$

20

## 【 0 0 6 4 】

式 ( 1 4 ) 及び ( 1 5 ) において、 $u$ 、 $v$  は、各々平均値がゼロ、分散が各々  $\sigma_u$ 、 $\sigma_v$  となるガウシアン乱数 ( gaussian random variable ) である。例えば、 $\sigma_u$  は 1 . 0 であり、 $\sigma_v$  は 0 . 2 である。

つまり、式 ( 1 4 ) は、状態遷移部 1 0 6 3 が、進行時間  $T$  中の仮想拍間隔  $b_i'$  ( テンポ ) に応じた楽譜位置の進行に一定範囲に分布するランダムな揺らぎを示す乱数  $u$  を加えて仮想楽譜位置  $k_i'$  を更新していることを表す。

30

他方、式 ( 1 5 ) は、状態遷移部 1 0 6 3 が、一定範囲に分布するランダムな揺らぎを示す乱数  $v$  を加えて仮想拍間隔  $b_i'$  を更新していることを表す。

このように、本実施形態は、粒子 i ごとに独立な状態遷移を行い、状態遷移において一定範囲内に分布する乱数を加えることで、演奏される音楽のテンポの揺らぎを考慮することができる。

状態遷移部 1 0 6 3 は、更新した粒子 i 毎の仮想楽譜位置  $k_i'$  を重み算出部 1 0 6 1 及び楽譜位置算出部 1 0 6 4 に出力し、更新した粒子 i 毎の仮想拍間隔  $b_i'$  を重み算出部 1 0 6 1 及びテンポ算出部 1 0 6 5 に出力する。

## 【 0 0 6 5 】

楽譜位置算出部 1 0 6 4 は、状態遷移部 1 0 6 3 から入力された粒子 i 毎の仮想楽譜位置  $k_i'$  に基づいて推定楽譜位置  $\mu$  を算出する。楽譜位置算出部 1 0 6 4 は、例えば、仮想楽譜位置  $k_i'$  の単純平均を  $\mu$  とする。本実施形態では、これに限られず、例えば楽譜位置算出部 1 0 6 4 は、仮想楽譜位置  $k_i'$  の偏差が所定の閾値  $u_{TH}$  よりも小さくなる粒子間の平均値を推定楽譜位置  $\mu$  と算出してよい。これにより、楽譜位置算出部 1 0 6 4 は、平均値を算出する際、稠密に分布している粒子だけを考慮し、極端に分散した粒子を除外できる。

40

楽譜位置算出部 1 0 6 4 は、粒子 i 毎の仮想楽譜位置  $k_i'$  と算出した推定楽譜位置  $\mu$  を信頼度算出部 1 0 7 に出力し、算出した推定楽譜位置  $\mu$  を楽譜位置出力部 1 0 8 に出力する。

## 【 0 0 6 6 】

50

信頼度算出部 107 は、楽譜位置算出部 1064 から入力された粒子  $i$  毎の仮想楽譜位置  $k'_i$  と算出した推定楽譜位置  $\mu$  に基づき、推定楽譜位置  $\mu$  の信頼度係数を算出する。信頼度算出部 107 は、例えば、次式 (16) を用いて最新の時刻  $t$  における信頼度係数  $s^2(t)$  を算出する。この信頼度係数  $s^2(t)$  は、粒子  $i$  間の仮想楽譜位置  $k'_i$  の分散であり、値が大きくなるほど推定楽譜位置の信頼性が低下することを示す。

【0067】

【数16】

$$s^2(t) = \sum_{i=1}^N (k'_i - \mu)^2 / N \quad (16)$$

10

【0068】

ここで、信頼度算出部 107 は、最新の時刻における信頼度係数  $s^2(t)$  と予め定めた遷移回数  $I$  (例えば 5 回) 前の時刻  $t - I \cdot T$  における信頼度係数  $s^2(t - I \cdot T)$  の差分 (以下、信頼度差分と呼ぶ) が予め定めた閾値  $inc \cdot I$  を超えるか否かを判断する。また、信頼度算出部 107 は、信頼度差分が予め定めた閾値  $dec \cdot I$  を下回るか否かを判断する。 $inc$  及び  $dec$  任意の実数であり、例えばともに 4 である。

信頼度算出部 107 は、信頼度差分が閾値  $inc \cdot I$  を超えると判断したとき、楽譜位置の出力を停止することを指示する楽譜位置出力停止信号を楽譜位置出力部 108 に出力する。信頼度算出部 107 は、信頼度差分が閾値  $dec \cdot I$  を下回ると判断したとき、楽譜位置の出力を開始することを指示する楽譜位置出力開始信号を楽譜位置出力部 108 に出力する。

20

【0069】

上述の例では、信頼度算出部 107 は、信頼度差分に基づき楽譜位置出力停止信号又は楽譜位置出力開始信号の出力を制御していたが、本実施形態ではこれに限らない。例えば、信頼度算出部 107 は、信頼度係数  $s^2(t)$  自体が予め定めた閾値  $s_0$  を超えるか否かを判断し、信頼度係数  $s^2(t)$  が閾値  $s_0$  を超えると判断したとき、楽譜位置出力停止信号を楽譜位置出力部 108 に出力し、信頼度係数  $s^2(t)$  が閾値  $s_0$  を下回ると判断したとき、楽譜位置出力開始信号を楽譜位置出力部 108 に出力してもよい。

30

信頼度算出部 107 は、上記の信頼度係数の代わりに推定楽譜位置  $\mu$  の推定誤差  $e(t)$  の絶対値  $|e(t)|$  もしくは二乗誤差  $e^2(t)$  を信頼度係数として算出し、算出した信頼度が予め定めた閾値  $e_0$  を超えるか否かを判断し、信頼度係数が閾値  $e_0$  を超えると判断したとき、楽譜位置出力停止信号を楽譜位置出力部 108 に出力し、信頼度係数が閾値  $e_0$  を下回ると判断したとき、楽譜位置出力開始信号を楽譜位置出力部 108 に出力してもよい。

ここで、推定誤差  $e(t)$  は、次式 (17) で表される。

【0070】

【数17】

$$e(t) = t + \Delta T - s(k'(t)) \quad (17)$$

40

【0071】

式 (17) において、 $s(k'(t))$  は、フレーム  $k'(t)$  における真の (ground truth) 時刻である。 $k'(t)$  は、推定楽譜位置  $\mu$  である。

【0072】

楽譜位置出力部 108 は、楽譜位置算出部 1064 から推定楽譜位置  $\mu$  が入力され、入力された推定楽譜位置  $\mu$  を出力する。

但し、楽譜位置出力部 108 は、信頼度算出部 107 から楽譜位置出力停止信号が入力

50

されたとき、推定楽譜位置  $\mu$  の出力を停止する。他方、楽譜位置出力部 108 は、信頼度算出部 107 から楽譜位置出力開始信号が入力されたとき、停止していた推定楽譜位置  $\mu$  の出力を再開する。

これにより、本実施形態は、信頼性が低い推定楽譜位置  $\mu$  の出力を回避することができる。

#### 【0073】

テンポ算出部 1065 は、状態遷移部 1063 から入力された粒子  $i$  毎の仮想拍間隔  $b_i$  に基づき推定拍間隔 を算出する。この推定拍間隔は、演奏されている音楽のテンポを示す尺度である。テンポ算出部 1065 は、例えば、仮想拍間隔  $b_i$  の単純平均をとす。本実施形態では、これに限られず、例えばテンポ算出部 1065 は、仮想拍間隔  $b_i$  の偏差が所定の閾値  $T_H$  よりも小さくなる粒子間の平均値を推定拍間隔 と算出してもよい。これにより、テンポ算出部 1065 は、平均値を算出する際に稠密に分布している粒子だけを考慮し、極端に分散した粒子を除外できる。

テンポ算出部 1065 は、推定拍間隔 をテンポ出力部 109 に出力する。

テンポ出力部 109 は、テンポ算出部 1065 から推定拍間隔 が入力され、入力された推定拍間隔 を楽譜位置推定装置 1 の外部に出力する。

#### 【0074】

ロボットは、例えば出力された推定楽譜位置  $\mu$  及び推定拍間隔 を歌声の合成や機構部の位置制御に用い、合成した歌声や機構部の動作を演奏されている音楽に同期させることができる。本実施形態の楽譜位置推定処理は、例えばロボットが演奏されている音楽に反応して歌声を合成したり、動作機構を制御する、等に応用することが期待されている。これらの応用において必然的に遅延が生じる問題を、ロボットが本実施形態により推定された楽譜位置を用いることで解決できる。

#### 【0075】

また、楽譜位置推定装置 1 は、当初、推定楽譜位置  $\mu$  と推定拍間隔 をともに出力するが、推定楽譜位置の信頼度増分が所定の閾値を越えると推定楽譜位置  $\mu$  の出力を停止する状態（リズムレベル (rhythm level) に遷移する。推定楽譜位置の信頼度増分が所定の閾値を下回ると推定楽譜位置  $\mu$  の出力を再開する状態（メロディレベル (melody level) に遷移する。これにより、本実施形態では信頼性が低い推定楽譜位置が出力されなくなるため、歌声の合成や機構部の位置制御等において、利用者の意図しない動作をすることを回避できる。

#### 【0076】

次に、本実施形態に係る楽譜位置推定処理の概要について図を用いて説明する。

図 2 は、本実施形態に係る楽譜位置推定処理の概要を示す概念図である。図 2 左列は観測 (I.)、中列は再標本化 (II.)、右列は状態遷移 (III.)、の過程を示す。このような過程を含む一連の処理は、粒子フィルタリングに含まれる処理である。

図 2 左列は、上段から下段へ順に、音響信号、楽譜情報、粒子の分布を示す。各段ともに横軸は時刻である。上段は、縦軸は振幅であり音響信号の波形を表す。中段は、五線上に音符を用いて楽譜情報を表す。下段は、円を用いて仮想楽譜位置毎の粒子の分布を表す。円の半径は、各粒子の重み係数の大きさを表す。

上段と中段に示されている四角の枠は、音響信号と楽譜情報を観測対象とする観測時間を示す。

従って、図 2 左列は、観測 (I.) の過程において、音響特徴量と楽譜特徴量に基づき、粒子毎に重み係数を算出することを示す。

#### 【0077】

図 2 中列は、中段に再標本化前の粒子の分布を、下段に再標本化後の粒子の分布を示す。ここで、再標本化前の粒子の分布は、同図左列下段に示されているものと同一である。再標本化前の粒子から下段に向かう破線の矢印と  $\times$  印は、その粒子が棄却されることを示す。再標本化前の粒子から再標本化後の粒子に向かう細い実線の矢印は、その粒子が維持されることを示す。再標本化前の粒子から再標本化後の粒子に向かう太い実線の矢印は、

その粒子が分割又は複製され、同一の仮想楽譜情報を有する粒子が生成されることを示す。

従って、図2中列は、再標本化において重み係数が小さい粒子が棄却され、重み係数が大きい粒子について分割又は複製されることにより棄却された数と同数の新たな粒子が生成されることを示す。

#### 【0078】

図2右列は、上段から下段へ順に、楽譜情報、遷移前の粒子の分布、遷移後の粒子の分布を示す。ここで、楽譜情報は、同図左列中段に示されているものと同一である。遷移前の粒子の分布は、同図中列下段に示されているものと同一である。中段において下に向かう矢印は、粒子の分布の中心が現在の楽譜位置であることを示す。下段において上に向かう矢印は、粒子の分布の中心が推定楽譜位置であることを示す。ここで、図2右列は、現在の楽譜位置と推定楽譜位置の間が、拍間隔であることを示す。中段の粒子から下段に向かう矢印は、その矢印が示す楽譜位置に粒子が遷移することを示す。

10

従って、図2右列は、状態遷移において粒子毎の仮想楽譜位置が更新され、更新後の仮想楽譜位置の分布に基づき楽譜位置が推定されることを示す。

#### 【0079】

図3は、本実施形態に係る楽譜位置推定処理の概要を示すフローチャートである。

(ステップS1) 楽譜位置推定装置1の各構成部は、処理に用いる変数やデータを初期設定する。

(ステップS2) 音響信号入力部101は、音響信号を入力し、音響信号の入力が継続するか否か判断する。音響信号入力部101が、音響信号の入力が継続すると判断したとき(ステップS2 Y)、ステップS3に進む。音響信号入力部101が、音響信号の入力が継続していないと判断したとき(ステップS2 N)、処理を終了する。

20

(ステップS3) 音響特徴量生成部102は、入力された音響信号に基づき特徴量を生成する。重み算出部1061は、生成した音響信号に基づく特徴量と楽譜情報に基づく特徴量から粒子毎の重み係数を算出する。(I. 観測)

#### 【0080】

(ステップS4) 再標本化部1062は、重み係数が所定の値より小さい粒子を棄却し、重み係数が所定の値より大きい粒子を分割又は複製する。(II. 再標本化)

(ステップS5) 状態遷移部1063は、粒子毎の仮想楽譜位置及び仮想拍間隔を更新する。楽譜位置算出部1064は、更新後の仮想楽譜位置に基づき推定楽譜位置を算出する。テンポ算出部1065は、更新後の仮想拍間隔に基づき推定拍間隔を算出する。(III. 状態遷移)

30

#### 【0081】

次に、本実施形態に係る楽譜位置推定処理の詳細について図を用いて説明する。

図4は、本実施形態に係る楽譜位置推定処理の一部を示すフローチャートである。

図4は、初期設定(ステップS1)、音響信号入力の継続判断(ステップS2)及び観測(ステップS3)の各過程を含んで構成される。

初期設定(ステップS1)は、ステップS101-S104を含んで構成される。

(ステップS101) 楽譜位置推定装置1の各構成部は、処理に用いる変数やデータを初期設定する。例えば、重み算出部1061は、粒子*i*毎の仮想拍間隔 $b_i'$ 、粒子*i*毎の仮想楽譜位置 $k_i'$ 、分散 $\sigma_u$ 、 $\sigma_v$ を初期設定する。

40

重み算出部1061は、粒子*i*毎の仮想拍間隔 $b_i'$ の初期値を、例えば各々テンポ*x*に対応する拍間隔 $60/x$ に定める。このテンポ*x*は、 $60 \text{ bpm}$  ( $\text{beat per minute}$ ; 1分あたりの拍数) から  $180 \text{ bpm}$  の間で均一に分布したものである。

重み算出部1061は、粒子*i*毎の仮想楽譜位置 $k_i'$ の初期値を、例えば、 $T/b_i'$ と定める。

重み算出部1061は、先行時間  $T$  を、例えば1秒と設定しておく。その後、ステップS102に進む。

(ステップS102) 楽譜情報入力部104は、楽譜情報記憶部103から、処理の対象

50



とする楽曲の楽譜情報  $n_k$  を読み出す。楽譜情報入力部 104 は、読み出した楽譜情報  $n_k$  を楽譜特徴量生成部 105 に出力する。その後、ステップ S103 に進む。

【0082】

(ステップ S103) 楽譜特徴量生成部 105 が備える周波数特性合成部 1051 は、楽譜情報入力部 104 から入力された楽譜情報  $n_k$  に基づいて楽譜特徴量として、例えば楽譜スペクトログラム  $X'_{f,k}$  を算出する。周波数特性合成部 1051 は、楽譜スペクトログラム  $X'_{f,k}$  を算出する際、例えば式(2)に表される harmonic GMM を用いる。

周波数特性合成部 1051 は、その楽曲を構成する楽譜情報  $n_k$  に全体について楽譜スペクトログラム  $X'_{f,k}$  を算出しておいてもよい。また、予め算出した楽譜スペクトログラム  $X'_{f,k}$  を楽譜情報記憶部 103 に記憶しておき、周波数特性合成部 1051 は、楽譜情報記憶部 103 から楽譜スペクトログラム  $X'_{f,k}$  を読み出してもよい。

周波数特性合成部 1051 は、生成または読み出した楽譜スペクトログラム  $X'_{f,k}$  を楽譜位置推定部 106 に出力する。その後、ステップ S104 に進む。

【0083】

(ステップ S104) 楽譜特徴量生成部 105 が備えるクロマベクトル生成部 1052 は、楽譜情報入力部 104 から入力された楽譜情報  $n_k$  に基づいて楽譜特徴量として、楽譜クロマベクトル  $c_k^s$  を生成する。

クロマベクトル生成部 1052 は、その楽曲を構成する楽譜情報  $n_k$  全てについて楽譜クロマベクトル  $c_k^s$  を生成しておいてもよい。また、予め生成した楽譜クロマベクトル  $c_k^s$  を楽譜情報記憶部 103 に記憶しておき、クロマベクトル生成部 1052 は、楽譜情報記憶部 103 から楽譜クロマベクトル  $c_k^s$  を読み出してもよい。

クロマベクトル生成部 1052 は、生成又は読み出した楽譜クロマベクトル  $c_k^s$  を楽譜位置推定部 106 に出力する。その後、ステップ S2 に進む。

【0084】

(ステップ S2) 音響信号入力部 101 は、音響信号を入力し、音響信号の入力が継続するか否かを判断する。音響信号入力部 101 が、音響信号の入力が継続すると判断したとき(ステップ S2 Y)、ステップ S301 に進む。音響信号入力部 101 が、音響信号の入力が継続していないと判断したとき(ステップ S2 N)、処理を終了する。

【0085】

観測(ステップ S3)の過程は、ステップ S301 - S307 を含んで構成される。

(ステップ S301) 音響信号入力部 101 は、アナログ音響信号を入力し、入力されたアナログ音響信号をデジタル信号に変換する。音響信号入力部 101 は、変換されたデジタル信号を音響特徴量生成部 102 に出力する。その後、ステップ S302 に進む。

(ステップ S302) 周波数特性分析部 1021 は、音響信号入力部 101 から入力されたデジタル音響信号を時間領域から周波数領域信号に変換する。周波数特性分析部 1021 は、変換された周波数領域信号の振幅を音響スペクトログラム  $X_{f,t}$  として算出し、算出した音響スペクトログラム  $X_{f,t}$  をクロマベクトル生成部 1022 及び楽譜位置推定部 106 に出力する。その後、ステップ S303 に進む。

(ステップ S303) クロマベクトル生成部 1022 は、周波数特性分析部 1021 から入力された音響スペクトログラム  $X_{f,t}$  から音響クロマベクトルを算出する。ここで、クロマベクトル生成部 1022 は、例えば、音響クロマベクトル  $c_t^a$  の音階  $n$  毎の要素  $c_t^a(n)$  を、式(1)を用いて算出する。

クロマベクトル生成部 1022 は、算出した音響クロマベクトル  $c_t^a$  を楽譜位置推定部 106 に出力する。その後、ステップ S304 に進む。

【0086】

(ステップ S304) クロマベクトル重み算出部 10611 は、クロマベクトル生成部 1022 から入力された音響クロマベクトル  $c_t^a$  とクロマベクトル生成部 1052 から入力された楽譜クロマベクトル  $c_{k,t}^s$  からクロマベクトル重み係数  $w^{c,h}$  を算出する。ここで、クロマベクトル重み算出部 10611 は、例えば式(4)及び(5)を用い、状態

10

20

30

40

50

遷移部 1063 から入力された粒子  $i$  毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  に基づき、粒子  $i$  毎のクロマベクトル重み係数  $w_i^{c,h}$  を算出する。

クロマベクトル重み算出部 10611 は、算出した粒子  $i$  毎のクロマベクトル重み係数  $w_i^{c,h}$  を粒子重み算出部 10614 に出力する。その後、ステップ S305 に進む。

【0087】

(ステップ S305) 周波数特性重み算出部 10612 は、周波数特性分析部 1021 から入力された音響スペクトログラム  $X_{f,t}$  と周波数特性合成部 1051 から入力された楽譜スペクトログラム  $X'_{f,k}$  から周波数特性重み係数  $w^{s,p}$  を算出する。ここで、周波数特性重み算出部 10612 は、例えば式 (5) - (8) を用い、状態遷移部 1063 から入力された粒子  $i$  毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  に基づき、粒子  $i$  毎の周波数特性重み係数  $w^{s,p}$  を算出する。

10

周波数特性重み算出部 10612 は、算出した粒子  $i$  毎の周波数特性重み係数  $w_i^{s,p}$  を粒子重み算出部 10614 に出力する。その後、ステップ S306 に進む。

【0088】

(ステップ S306) テンポ重み算出部 10613 は、周波数特性分析部 1021 から入力された音響スペクトログラム  $X_{f,t}$  からテンポ重み係数  $w^t$  を算出する。ここで、テンポ重み算出部 10613 は、例えば式 (9) を用いて状態遷移部 1063 から入力された粒子  $i$  毎の仮想拍間隔  $b_i'$  に基づき、粒子  $i$  毎のテンポ重み係数  $w^t$  を算出する。

テンポ重み算出部 10613 は、算出した粒子  $i$  毎のテンポ重み係数  $w_i^t$  を粒子重み算出部 10614 に出力する。その後、ステップ S307 に進む。

20

【0089】

(ステップ S307) 粒子重み算出部 10614 は、クロマベクトル重み算出部 10611 から入力されたクロマベクトル重み係数  $w_i^{c,h}$ 、周波数特性重み算出部 10612 から入力された周波数特性重み係数  $w_i^{s,p}$ 、及びテンポ重み算出部 10613 から入力されたテンポ重み係数  $w_i^t$  から粒子  $i$  毎の重み係数  $w_i$  を算出する。ここで、粒子重み算出部 10614 は、重み係数  $w_i$  を算出する際、例えば式 (10) を用いる。

粒子重み算出部 10614 は、算出した粒子  $i$  毎の仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  及び粒子重み係数  $w_i$  を再標本化部 1062 に出力する。その後、ステップ S411 に進む。

【0090】

30

図 5 は、本実施形態に係る楽譜位置推定処理の他の一部を示すフローチャートである。図 5 は、再標本化 (ステップ S4) の過程を含んで構成される。再標本化 (ステップ S4) は、棄却 (ステップ S41)、分割 (ステップ S42) 及び複製 (ステップ S43) の各過程を含んで構成される。

棄却 (ステップ S41) の過程は、ステップ S411 - S414 を含んで構成される。(ステップ S411) 再標本化部 1062 は、粒子重み算出部 10614 から粒子  $i$  毎の仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  及び粒子重み係数  $w_i$  を入力する。

再標本化部 1062 は、全ての粒子  $i$  についてステップ S412 - S414 の処理を実行したか否か判断する。再標本化部 1062 は、全ての粒子  $i$  についてステップ S412 - S414 の処理を実行したと判断したとき (ステップ S411 Y)、ステップ S421 に進む。再標本化部 1062 は、全ての粒子  $i$  についてステップ S412 - S414 の処理を実行していないと判断したとき (ステップ S411 N)、ステップ S412 に進む。

40

【0091】

(ステップ S412) 再標本化部 1062 は、例えば、式 (11) を用いて、粒子  $i$  毎の粒子重み係数  $w_i$  から規格化粒子重み係数  $p_i$  を算出する。再標本化部 1062 は、規格化粒子重み係数  $p_i$  が予め設定された閾値  $p_{t,h}$  よりも小さいか否か判断する。再標本化部 1062 は、規格化粒子重み係数  $p_i$  が閾値  $p_{t,h}$  よりも小さいと判断した場合 (ステップ S411 Y)、ステップ S413 に進む。再標本化部 1062 は、規格化粒子重み係数  $p_i$  が閾値  $p_{t,h}$  よりも大きい又は等しいと判断した場合 (ステップ S411 N)

50

、対応する粒子  $i$  を維持し、ステップ S 4 1 1 に進む。

(ステップ S 4 1 3) 再標本化部 1 0 6 2 は、規格化粒子重み係数  $p_i$  が閾値  $p_{th}$  よりも小さいと判断された粒子  $i$  を棄却する。その後、ステップ S 4 1 4 に進む。

(ステップ S 4 1 4) 再標本化部 1 0 6 2 は、棄却された粒子数を 1 個増加させ、棄却された粒子数をカウントする。その後、ステップ S 4 1 1 に進む。

#### 【 0 0 9 2 】

分割 (ステップ S 4 2) の過程は、ステップ S 4 2 1 - S 4 2 6 を含んで構成される。

(ステップ S 4 2 1) 再標本化部 1 0 6 2 は、棄却されずに残った粒子  $i$  についてステップ S 4 2 2 - S 4 2 6 の処理を実行したか否か判断する。再標本化部 1 0 6 2 は、全ての残った粒子  $i$  についてステップ S 4 2 2 - S 4 2 6 の処理を実行したと判断したとき (ステップ S 4 2 1 Y)、ステップ S 4 3 1 に進む。再標本化部 1 0 6 2 は、全ての粒子  $i$  についてステップ S 4 2 2 - S 4 2 6 の処理を実行していないと判断したとき (ステップ S 4 2 1 N)、ステップ S 4 2 2 に進む。

(ステップ S 4 2 2) 再標本化部 1 0 6 2 は、例えば、式 (1 2) を用いて、粒子  $i$  毎の規格化粒子重み係数  $p_i$  から再規格化粒子重み係数  $q_i$  を算出する。再標本化部 1 0 6 2 は、再規格化粒子重み係数  $q_i$  が 1 よりも大きい粒子について再規格化整数  $n_{q_i}$  と再規格化小数  $s_{q_i}$  を算出する。その後、ステップ S 4 2 3 に進む。

#### 【 0 0 9 3 】

(ステップ S 4 2 3) 再標本化部 1 0 6 2 は、再規格化整数  $n_{q_i}$  が 2 又は 2 より多いか否か判断する。再標本化部 1 0 6 2 は、再規格化整数  $n_{q_i}$  が 2 又は 2 より多いと判断したとき (ステップ S 4 2 3 Y)、ステップ S 4 2 4 に進む。再標本化部 1 0 6 2 が、再規格化整数  $n_{q_i}$  が 1 であると判断したとき (ステップ S 4 2 3 N)、ステップ S 4 2 1 に進む。

(ステップ S 4 2 4) 再標本化部 1 0 6 2 は、粒子  $i$  を分割して、新たに  $n_{q_i} - 1$  個の粒子  $j$  ( $j \neq i$ ) を生成する。即ち、再標本化部 1 0 6 2 は、新たな粒子  $j$  に対応する仮想楽譜位置  $k_j'$  及び仮想拍間隔  $b_j'$  を分割前の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  と定める。その後、ステップ S 4 2 5 に進む。

#### 【 0 0 9 4 】

(ステップ S 4 2 5) 再標本化部 1 0 6 2 は、不足粒子数  $N_s$  (初期値は棄却された粒子数) から新たに生成した粒子数  $n_{q_i} - 1$  を差し引いて不足粒子数をカウントする。その後、ステップ S 4 2 6 に進む。

(ステップ S 4 2 6) 再標本化部 1 0 6 2 は、不足粒子数  $N_s$  がゼロに達したか否か判断する。再標本化部 1 0 6 2 は、不足粒子数  $N_s$  がゼロに達したと判断したとき (ステップ S 4 2 6 Y)、再標本化に関する処理を終了し、粒子  $i$  毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  を状態遷移部 1 0 6 3 に出力する。その後、ステップ S 5 0 1 に進む。不足粒子数  $N_s$  がゼロに達していないと判断したとき (ステップ S 4 2 6 N)、ステップ S 4 2 1 に進む。

#### 【 0 0 9 5 】

複製 (ステップ S 4 3) の過程は、ステップ S 4 3 1 - S 4 3 4 を含んで構成される。

(ステップ S 4 3 1) 再標本化部 1 0 6 2 は、棄却されずに残った粒子  $i$  についてステップ S 4 3 2 - S 4 3 3 の処理を実行したか否か判断する。再標本化部 1 0 6 2 は、全ての残った粒子  $i$  についてステップ S 4 3 2 - S 4 3 3 の処理を実行したと判断したとき (ステップ S 4 3 1 Y)、ステップ S 4 3 4 に進む。再標本化部 1 0 6 2 は、全ての粒子  $i$  についてステップ S 4 3 2 - S 4 3 3 の処理を実行していないと判断したとき (ステップ S 4 3 1 N)、ステップ S 4 3 2 に進む。

(ステップ S 4 3 2) 再標本化部 1 0 6 2 は、例えば式 (1 3) を用いて再規格化小数  $s_{q_i}$  に比例する複製確率  $s'_{q_i}$  を算出する。その後、ステップ S 4 3 3 に進む。

(ステップ S 4 3 3) 再標本化部 1 0 6 2 は、複製確率  $s'_{q_i}$  に基づいて新たに複製される粒子の総数が不足粒子数  $N_s$  となるように、複製対象の粒子  $i$  を決定する。その後、ステップ S 4 3 1 に進む。

(ステップS434)再標本化部1062は、定めた粒子各々について、新たに1個の粒子( $j$   $i$ )を複製する。即ち、再標本化部1062は、新たな粒子 $j$ に対応する仮想楽譜位置 $k_j'$ 及び仮想拍間隔 $b_j'$ を複製前の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ と定める。

再標本化部1062は、粒子 $i$ 毎の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ を状態遷移部1063に出力する。その後、ステップS501に進む。

#### 【0096】

図6は、本実施形態に係る楽譜位置推定処理の残りの一部を示すフローチャートである。図6は、状態遷移(ステップS5)及び推定値出力(ステップS6)の過程を含んで構成される。状態遷移(ステップS5)の過程は、ステップS501-S504を含んで構成される。

10

(ステップS501)状態遷移部1063は、再標本化部1062から粒子 $i$ 毎の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ を入力する。状態遷移部1063は、進行時間 $T$ 中に仮想拍間隔 $b_i'$ (テンポ)で進行した楽譜位置に乱数 $u$ を加えて仮想楽譜位置 $k_i'$ を更新する。状態遷移部1063は、仮想楽譜位置 $k_i'$ を更新する際、例えば式(14)を用いる。その後、ステップS502に進む。

(ステップS502)状態遷移部1063は、乱数 $v$ を加えて仮想拍間隔 $b_i'$ を更新する。状態遷移部1063は、仮想拍間隔 $b_i'$ を更新する際、例えば式(15)を用いる。状態遷移部1063は、更新した粒子 $i$ 毎の仮想楽譜位置 $k_i'$ 及び仮想拍間隔 $b_i'$ を重み算出部1061に出力する。状態遷移部1063は、更新した粒子 $i$ 毎の仮想楽譜位置 $k_i'$ を楽譜位置算出部1064に出力する。状態遷移部1063は、更新した粒子 $i$ 毎の仮想拍間隔 $b_i'$ をテンポ算出部1065に出力する。その後、ステップS503に進む。

20

#### 【0097】

(ステップS503)楽譜位置算出部1064は、状態遷移部1063から入力された粒子 $i$ 毎の仮想楽譜位置 $k_i'$ に基づいて推定楽譜位置 $\mu$ を算出する。楽譜位置算出部1064は、例えば、仮想楽譜位置 $k_i'$ の単純平均を $\mu$ とする。楽譜位置算出部1064は、粒子 $i$ 毎の仮想楽譜位置 $k_i'$ と算出した推定楽譜位置 $\mu$ を信頼度算出部107に出力する。楽譜位置算出部1064は、算出した推定楽譜位置 $\mu$ を楽譜位置出力部108に出力する。その後、ステップS503に進む。

30

(ステップS504)テンポ算出部1065は、状態遷移部1063から入力された粒子 $i$ 毎の仮想拍間隔 $b_i'$ に基づいて推定拍間隔を算出する。テンポ算出部1065は、例えば、仮想拍間隔 $b_i'$ の単純平均をとする。

テンポ算出部1065は、推定拍間隔をテンポ出力部109に出力する。その後、ステップS601に進む。

#### 【0098】

推定値出力(ステップS6)の過程は、ステップS601-S606を含んで構成される。

(ステップS601)信頼度算出部107は、楽譜位置算出部1064から入力された粒子 $i$ 毎の仮想楽譜位置 $k_i'$ に基づき推定楽譜位置 $\mu$ を算出する。信頼度算出部107は、仮想楽譜位置 $k_i'$ と推定楽譜位置 $\mu$ に基づき、推定楽譜位置 $\mu$ の信頼度係数 $s^2(t)$ を算出する。信頼度算出部107は、信頼度係数 $s^2(t)$ を算出する際、例えば、式(16)を用いる。その後、ステップS602に進む。

40

(ステップS602)信頼度算出部107は、信頼度差分が予め定めた閾値 $i^{nc}I$ を超えるか否かを判断する。信頼度算出部107は、信頼度差分が予め定めた閾値 $i^{nc}I$ を超えると判断したとき(ステップS602 Y)、ステップS603に進む。信頼度算出部107は、信頼度差分が予め定めた閾値 $i^{nc}I$ を超えないと判断したとき(ステップS602 N)、ステップS604に進む。

(ステップS603)信頼度算出部107は、楽譜位置出力停止信号を楽譜位置出力部108に出力する。ここで、楽譜位置出力部108は、楽譜位置算出部1064から推定楽

50

譜位置  $\mu$  が入力され、入力された推定楽譜位置  $\mu$  を出力する。楽譜位置出力部 108 は、信頼度算出部 107 から楽譜位置出力停止信号が入力されたとき、推定楽譜位置  $\mu$  の出力を停止する。その後、ステップ S604 に進む。

#### 【0099】

(ステップ S604) 信頼度算出部 107 は、信頼度算出部 107 は、信頼度差分が予め定めた閾値 -  $d^{ec}I$  を下回るか否かを判断する。信頼度算出部 107 は、信頼度差分が予め定めた閾値 -  $d^{ec}I$  を下回ると判断したとき (ステップ S604 Y)、ステップ S605 に進む。信頼度算出部 107 は、信頼度差分が予め定め閾値 -  $d^{ec}I$  を下回らないと判断したとき (ステップ S604 N)、ステップ S606 に進む。

(ステップ S605) 信頼度算出部 107 は、楽譜位置出力開始信号を楽譜位置出力部 108 に出力する。楽譜位置出力部 108 は、信頼度算出部 107 から楽譜位置出力開始信号が入力されたとき、停止していた推定楽譜位置  $\mu$  の出力を再開する。その後、ステップ S606 に進む。

(ステップ S606) テンポ出力部 109 は、テンポ算出部 1065 から推定拍間隔  $\tau$  が入力され、入力された推定拍間隔  $\tau$  を出力する。その後、ステップ S7 に進む。

#### 【0100】

(ステップ S7) 重み算出部 1061 は、状態遷移部 1063 から粒子  $i$  毎の仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  が入力される。クロマベクトル重み算出部 10611 は、観測時間の始点  $t - L$  と終点  $t$  を、各々先行時間  $T$  だけ後の時刻にシフトさせて観測時間を変更する。

重み算出部 1061 は、入力された仮想楽譜位置  $k_i'$ 、仮想拍間隔  $b_i'$  及び観測時間を、クロマベクトル重み算出部 10611 及び周波数特性重み算出部 10612 に出力する。重み算出部 1061 は、入力された仮想拍間隔  $b_i'$  及び観測時間を、テンポ重み算出部 10613 に出力する。重み算出部 1061 は、仮想楽譜位置  $k_i'$  及び仮想拍間隔  $b_i'$  を、粒子重み算出部 10614 に出力する。その後、ステップ S2 に進む。

#### 【0101】

次に、本実施形態に係る楽譜位置推定装置 1 によって推定された楽譜位置の例について図を用いて説明する。図 7 - 14 は、特に断らない限り粒子数  $N$  が 300 の場合における楽譜位置推定装置 1 の動作結果を示す。

図 7 は、本実施形態に係る楽譜位置の推定誤差 (estimation error、prediction error) の一例を示す図である。図 7 において、縦軸は楽譜位置の推定誤差 (秒) を示し、横軸は時刻 (秒) を示す。

図 7 によれば、時刻 0 - 10 秒又は 20 - 30 秒付近では、推定誤差は概ね  $\pm 3$  秒未満である。時刻 35 - 40 秒付近での推定誤差が 15 秒前後と、より大きくなる。さらに、時刻 41 - 50 秒でも推定誤差は概ね  $\pm 3$  秒未満となる。

#### 【0102】

図 8 は、本実施形態に係る信頼度係数の一例を示す図である。図 8 において、縦軸は信頼度係数を示し、横軸は時刻 (秒) を示す。図 8 によれば、時刻 0 - 10 秒又は 20 - 30 秒付近では、信頼度係数が概ね 20 未満である。図 7 の結果を考慮すると、粒子が正しい楽譜位置に収束していることを示す。他方、信頼度係数が時刻 35 - 40 秒付近で約 60 から 100 へと増加する。この結果は、信頼度係数の増加が推定誤差の累積によって生じることを示す。また、図 8 によれば、時刻 41 - 50 秒付近で信頼度係数が低下する。この低下は、粒子が正しい楽譜位置に収束することを意味する。

従って、本実施形態では、信頼度算出部 107 が、信頼度差分が所定の閾値以上であることを検知することで、顕著な推定誤差をもたらす推定楽譜位置の出力を停止することができる。また、信頼度算出部 107 が、信頼度差分が所定の閾値以下であることを検知することで、推定誤差が小さい推定楽譜位置の出力を開始することができる。これにより、推定誤差が著しい推定楽譜位置の出力を回避できることが裏付けられる。

#### 【0103】

図 9 は、本実施形態に係る推定誤差が生じる割合の一例を示す図である。図 9 において

10

20

30

40

50

、縦軸は、推定誤差の絶対値が1秒未満となる割合（以下、低推定誤差率（rate of low prediction error）と呼ぶ）であり、横軸は、テンポ（bpm）である。図9において、 $\square$ 及び $\times$ は、演奏された楽曲のテンポ毎の低推定誤差率である。 $\square$ は、ドラム音を含む楽曲を示す。 $\times$ は、ドラム音を含まない楽曲を示す。

図9は、テンポが110bpm以下の楽曲のほとんどでは、低推定誤差率は0.9よりも高く、楽譜位置が精度よく推定されていることを示す。他方、図9は、テンポが120bpm以上の楽曲では、低推定誤差率は0.1よりも低く、楽譜位置が良好に推定されていないことを示す。その理由は、120bpm以上のテンポでは、拍間隔 $b$ が0.5秒以下となるため、先行時間 $T$ （この例では1秒）内にテンポ重み係数 $w^t$ （式(9)参照）のピークが2個以上存在するからである。つまり、一意にテンポ重み係数 $w^t$ が定まらないことが推定誤差を増加させる要因と考えられる。

10

#### 【0104】

また、図9は、テンポが110bpm以下において、ドラム音を含む楽曲よりもドラム音を含まない楽曲のほうが、楽譜位置が精度よく推定されていることを示す。これは、ドラム音の非調和性（inharmonic nature）がクロマベクトル重み係数 $w_i^{ch}$ 、周波数特性重み算出部10612から入力された周波数特性重み係数 $w_i^{sp}$ を低下させる有力な原因となるからである。

なお、上述の低推定誤差率とは異なり、0.5秒未満又は2秒未満の推定誤差をとる割合を低推定誤差率と定めた場合でも、本実施形態では図9と同様な結果が得られる。

#### 【0105】

20

図10は、本実施形態に係る推定誤差の一例を示す図である。図10において、縦軸は、推定誤差であり、横軸は、テンポ（bpm）である。図10において、 $\square$ 及び $\times$ は、演奏された楽曲のテンポ毎の平均推定誤差である。 $\square$ は、ドラム音を含む楽曲を示す。 $\times$ は、ドラム音を含まない楽曲を示す。また、 $\square$ 及び $\times$ の上下に伸びる棒線は、推定誤差の標準偏差を示す。

#### 【0106】

図10は、テンポが110bpm以下の楽曲では、平均推定誤差が-10秒から0秒の範囲であり、テンポが120bpm以上の楽曲では、平均推定誤差が10秒から20秒に分布している。また、図10は、テンポが110bpm以下の楽曲のほとんどでは、標準偏差は7秒以内にとどまるのに対し、テンポが120bpm以上の楽曲では、標準偏差は10秒前後となる。即ち、図10は、テンポが110bpm以下の楽曲では、テンポが120bpm以上の楽曲よりも、楽譜位置が精度よく推定されることを示す。

30

また、図10は、テンポが110bpm以下において、ドラム音を含まない楽曲よりもドラム音を含む楽曲のほうが、標準偏差が大きいことを示す。この点において、ドラム音を含む楽曲よりもドラム音を含まない楽曲のほうが、楽譜位置が精度よく推定されることを裏付ける。

#### 【0107】

図11は、本実施形態に係る推定誤差が生じる割合の他の例を示す図である。

この例は、楽譜位置出力部108が推定楽譜位置を出力する場合（メロディレベル）のみを考慮して算出した低推定誤差率を示す。

40

図11において、縦軸は、低推定誤差率であり、横軸は、テンポ（bpm）である。図11において、 $\square$ 及び $\times$ は、演奏された楽曲のテンポ毎の低推定誤差率である。 $\square$ は、ドラム音を含む楽曲を示す。 $\times$ は、ドラム音を含まない楽曲を示す。

#### 【0108】

図11は、テンポが110bpm以下の楽曲のほとんどでは、低推定誤差率は0.9よりも高く、楽譜位置が精度よく推定されていることを示す。他方、図11は、テンポが120bpm以上の楽曲では、低推定誤差率は0.2から0.8の間を分布し、楽譜位置が良好に推定されていないことを示す。しかし、図11が示す低推定誤差率は、図9が示す低推定誤差率よりも高い。この結果は、信頼度算出部107が算出する信頼度に基づいて、楽譜位置出力部108からの推定楽譜位置の出力の有無を制御する構成が、楽譜位置の

50

推定精度を向上させることを示す。

【0109】

図12は、本実施形態に係る推定誤差の他の例を示す図である。

この例も、楽譜位置出力部108が推定楽譜位置を出力する場合（メロディレベル）のみを考慮して算出した低推定誤差率を示す。

図12において、縦軸は、推定誤差であり、横軸は、テンポ（bpm）である。図12において、 $\Delta$ 及び $\times$ は、演奏された楽曲のテンポ毎の平均推定誤差である。 $\Delta$ は、ドラム音を含む楽曲を示す。 $\times$ は、ドラム音を含まない楽曲を示す。また、 $\Delta$ 及び $\times$ の上下に伸びる棒線は、推定誤差の標準偏差を示す。

【0110】

図12は、テンポが110bpm以下の楽曲では、平均推定誤差が-2秒から2秒の範囲であり、テンポが120bpm以上のほとんどの楽曲では、平均推定誤差が10秒から15秒に分布している。また、図12は、テンポが110bpm以下の楽曲のほとんどでは、標準偏差は1秒以内にとどまるのに対し、テンポが120bpm以上の楽曲では、標準偏差は7~10秒となる。図10に示す結果と比較すると、図12に示す平均推定誤差、標準偏差ともに低くなる。

この結果も、信頼度算出部107が算出する信頼度に基づいて、楽譜位置出力部108からの推定楽譜位置の出力の有無を制御する構成が、楽譜位置の推定精度を向上させることを示す。

【0111】

図13は、本実施形態に係る推定誤差と粒子数の関係の一例を示す図である。

図13において、縦軸は推定誤差を、横軸は粒子数を示す。 $\Delta$ 、 $\times$ 及び $\square$ は、テンポが98、122、及び200bpmの楽曲について、粒子数毎の平均推定誤差を示す。 $\Delta$ 、 $\times$ 及び $\square$ の上下に伸びる棒線は、粒子数毎の標準偏差を示す。

【0112】

図13によれば、テンポが98bpmの楽曲について、平均推定誤差は、粒子数150について-5秒であるのに対し、粒子数300以上について-2~-3秒に収束する。標準偏差は、粒子数150について上下4秒であるのに対し、粒子数300以上について上下1~2秒に収束する。

テンポが122bpmの楽曲について、平均推定誤差は、粒子数150について1.4秒であるのに対し、粒子数300以上について3~5秒に収束する。標準偏差は、粒子数150について $\pm 8$ 秒であるのに対し、粒子数300以上について $\pm 3$ 秒に収束する。

これに対し、テンポが200bpmの楽曲について、粒子数に関わらず平均推定誤差は1.5秒前後である。標準偏差は粒子数に関わらず $\pm 8$ 秒である。

即ち、図13の結果は、テンポが約120bpm以下の楽曲について、粒子数が300あれば楽譜位置を精度よく推定できることを示す。反面、テンポが高い楽曲については楽譜位置を精度よく推定できないことを示す。

【0113】

図14は、本実施形態に係る精度（precision）と再現率（recall）の一例を示す図である。

図14において、縦軸は精度及び再現率を示し、横軸はテンポ（bpm）を示す。

ここで、精度  $p$  は  $x^t / y$  と、再現率  $r$  は  $x^t / z$  である。但し、 $x^t$  は、推定誤差  $e(t)$  の絶対値が1秒を越える場合であって、リズムレベル（楽譜位置出力部108から推定楽譜位置を出力する場合）における出力の個数である。 $y$  は、リズムレベルにおける全ての出力の個数である。 $z$  は、推定誤差  $e(t)$  の絶対値が1秒を超える場合の出力の個数である。

【0114】

図14において、 $\times$  は演奏された楽曲のテンポ毎の精度を、 $\Delta$  は演奏された楽曲の再現率を示す。図14によれば、ほとんどの楽曲（特にテンポが120bpm以上の全ての楽曲）について精度が概ね1である。これは、信頼度算出部107において信頼度に基づい

10

20

30

40

50

て推定楽譜位置の出力を停止する場合は、概ね推定誤差が著しいことを示す。図14によれば、再現率の平均は0.43である。ここで、テンポが110bpm以下の楽曲について、2曲を除いて再現率は0.05未満となり、テンポが120bpm以上の楽曲について、再現率は0.2~0.8の間で分布する。テンポが110bpm以下の楽曲について、そもそも推定誤差 $e(t)$ の絶対値が1秒を超える場合が少ないため、全体として推定精度の低下させずに済む。

#### 【0115】

上述の結果からも、本実施形態は、従来困難だった拍構造が不明確な楽曲（例えば、ドラム音を含まない楽曲）であっても演奏されている楽譜位置を精度よく推定できる。また、上述の結果において、テンポが速い楽曲について推定誤差が増加する傾向が認められた。その場合でも、先行時間  $T$  をより短くすれば、テンポが速い楽曲でも楽譜位置を精度よく推定できることが示唆される。

このように、本実施形態は、入力された音響信号の特徴量と楽譜情報の特徴量に基づいて重み係数を算出し、その重み係数に対応する仮想楽譜位置及び仮想テンポを用いて楽譜位置を推定する。そのため、本実施形態によれば演奏された音楽に対する楽譜位置を精度よく推定することが可能になる。

#### 【0116】

（第2の実施形態）

以下、図面を参照しながら本発明の第2の実施形態について説明する。

図15は、本実施形態に係る楽譜位置推定装置2の構成を示す概略図である。

楽譜位置推定装置2は、楽譜位置推定装置1の楽譜位置推定部106及び信頼度算出部107の代わりに楽譜位置推定部206及び信頼度算出部207を備える。

楽譜位置推定部206は、楽譜位置推定部106の重み算出部1061の代わりに重み算出部2061を備える。重み算出部2061は、クロマベクトル重み算出部10611、周波数特性重み算出部10612、及びテンポ重み算出部10613を備える点で重み算出部1061と同様である。

また、楽譜位置推定装置2のその他の機能及び構成は第1の実施形態に係る楽譜位置推定装置1と同様である。以下、本実施形態について、第1の実施形態との差異点を主に説明する。

#### 【0117】

信頼度算出部207は、例えば、上述の方式の何れかによって算出した信頼度係数を楽譜位置推定部206に出力する点が、信頼度算出部107と異なる。信頼度算出部207のその他の機能及び構成は信頼度算出部107と同様である。

なお、信頼度算出部207は、楽譜位置推定部206に出力する信頼度係数として、楽譜位置出力停止信号及び楽譜位置出力開始信号の出力の要否を判断するために用いる信頼度係数とは異なる方式で算出してもよい。例えば、信頼度算出部207は、後述するDTW (dynamic time warping; 動的時間収縮) 法を用いて距離を算出し、算出した距離を信頼度係数として楽譜位置推定部206に出力してもよい。後述するように、信頼度算出部207は、このDTW法を用いて音響特徴量と楽譜特徴量に基づき距離を算出する過程で、音響特徴量の区間と合致する楽譜特徴量の区間を定めることができる。

本実施形態では、この距離を算出する構成を重み算出部2061に備え、この距離を算出してもよい。

#### 【0118】

重み算出部2061は、信頼度算出部207から入力された信頼度係数に基づき観測時間を変更するかどうか判断する。ここで、観測時間を変更する処理について説明する。

図16は、本実施形態に係る観測時間を変更する処理の概要を示す概念図である。

図16上段において、縦軸は振幅、横軸は時刻である、曲線は入力音響信号の波形の一例を表す。図16上段左側は、楽譜位置推定装置1の重み算出部1061と同様に観測時間をシフトする処理 (A.シフト) を示す。実線の四角形は、現時刻の観測時間を示し、破

10

20

30

40

50



線の四角形は次時刻の観測時間を示す。即ち、図16上段左側は、観測時間長 $L$ を変更せずに先行時間 $T$ だけ後の時刻に観測時間が変更することを示す。

図16上段右側は、観測時間を延長させる処理(B.延長)を示す。実線の四角形は、現時刻の観測時間を示し、破線の四角形は次時刻の観測時間を示す。即ち、図16上段右側は、観測開始時刻を固定し観測終了時刻を先行時間 $T$ だけ後の時刻に変更することにより、観測時間長が $L$ から $L + T$ に延長することを示す。

【0119】

図16中段及び下段は、観測時間を現時刻までに入力された音響信号全体に延長する処理(C.全区間に延長)を示す。図16中段及び下段ともに、横軸は時刻を示す。図16中段において、縦軸は振幅を示し、曲線は入力音響信号の波形の一例を表す。図16下段は、楽譜情報を五線上に音階とその長さを示す音符を用いて示す。

10

図16中段右側の実線の四角形は、現時刻の観測時間を示し、図16下段右側の実線の四角形は同図中段の実線の四角形と対応する区間の楽譜情報を示す。図16中段の破線の四角形は、次時刻の観測時間を示し、図16下段の破線の四角形は同図中段の破線の四角形と対応する区間の楽譜情報を示す。即ち、図16中段及び下段は、観測時間長 $L + T$ の観測時間を現時刻までに入力された音響信号全体 $L^{a11}$ に延長することを示す。

【0120】

次に本実施形態に係る重み算出部2061が観測時間を変更する処理について説明する。

図17は、本実施形態に係る観測時間を変更する処理を示すフローチャートである。

20

重み算出部2061は、以下に説明する処理を、図6のステップS7の代わりに行う。(ステップS71)重み算出部2061は、信頼度算出部207から信頼度係数が入力される。その後、ステップS72に進む。

(ステップS72)重み算出部2061は、入力された信頼度係数が予め設定した閾値よりも大きい(即ち、推定楽譜時刻の信頼性が低い)か否か判断する。重み算出部2061は、信頼度が予め設定した閾値よりも小さいと判断したとき(ステップS72 N)、ステップS73に進む。重み算出部2061は、信頼度係数が予め設定した閾値よりも大きい又は等しいと判断したとき(ステップS72 Y)、ステップS74に進む。

(ステップS73)重み算出部2061は、観測時間を、観測時間長 $L$ を変更せずに先行時間 $T$ だけシフトさせる(A.シフト)。その後、図4のステップS2に進む。

30

(ステップS74)重み算出部2061は、誤り回数を1加算してカウントする。誤り回数とは、楽譜位置の推定が誤った回数であり、初期値はゼロである。その後、ステップS75に進む。

【0121】

(ステップS75)重み算出部2061は、誤り回数が予め設定した閾値(例えば、5)よりも大きいか否か判断する。重み算出部2061は、誤り回数が予め設定した閾値よりも大きいと判断したとき(ステップS75 Y)、ステップS77に進む。重み算出部2061は、誤り回数が予め設定した閾値よりも小さい又は等しいと判断したとき(ステップS75 N)、ステップS76に進む。

(ステップS76)重み算出部2061は、観測時間の開始時刻を変更せず、終了時刻を先行時間 $T$ だけ延長することで観測時間長を延長する(B.延長)。誤り回数が予め設定した閾値(例えば、5)よりも大きいか否か判断する。その後、図4のステップS2に進む。

40

(ステップS77)重み算出部2061は、観測時間を現時刻までに入力された音響信号全体に延長する(C.全区間に延長)。その後、図4のステップS2に進む。

【0122】

これにより、楽譜位置の推定誤差が著しい場合でも、本実施形態が音響信号(又は音響特徴量)と楽譜情報(又は楽譜特徴量)の観測時間を延長して、両特徴量が合致する部分が見出される可能性が高くなる。これにより、本実施形態に係る楽譜位置推定装置2は、推定楽譜位置の精度が低下した状態から精度がより向上した状態に回復することができる

50

。ひいては、本実施形態は、推定楽譜位置の信頼性が低く出力されない状態から、信頼性の低下を回避して推定楽譜位置の出力を再開することができる。

【 0 1 2 3 】

次に、上述のDTW法を用いて信頼度係数を算出する処理について説明する。

DTW法では、重み算出部2061又は信頼度算出部207（以下、重み算出部2061等と呼ぶ）は、距離の算出対象とする音響特徴量の一部の区間と楽譜特徴量の一部の区間をマッチングさせ、全体の距離（累積距離）を最小化する音響特徴量の区間と楽譜特徴量の区間の組を決定する。即ち、重み算出部2061等は、この最小化された累積距離を信頼度係数と定める。以下の例では、このような音響特徴量の区間の始点を音響始点、終点を音響終点と呼び、このような楽譜特徴量区間の始点を楽譜始点、終点を楽譜終点と呼ぶ。

10

【 0 1 2 4 】

図18は、本実施形態に係る楽譜位置推定装置2が行うDTW法を示すフローチャートである。

（ステップS711）重み算出部2061等は、楽譜終点 $n_2$ を、例えば楽譜特徴量の時系列データ冒頭のフレームを含むサブ系列に初期設定する。

楽譜終点 $n_2$ は、ここでは楽譜特徴量の時系列データにおいて累積距離 $D_{n_2}$ を算出するための区間の終点、つまり、その区間の楽譜特徴量を含むサブ系列（*subsequence*）である。サブ系列とは、楽譜特徴量の時系列データ及び音響特徴量の時系列データについて両者をマッチングすることにより両者が合致する区間を探索する単位となる小区間である。

20

本実施形態では、観測時間が複数個のサブ系列からなり、1個のサブ系列に整数個の楽譜フレームが含まれるように、サブ系列の長さ（例えば、2楽譜フレーム、1/24秒に相当）を予め定めておく。その後、ステップS712に進む。

【 0 1 2 5 】

（ステップS712）重み算出部2061等は、楽譜終点 $n_2$ が楽譜特徴量の時系列データの最初のサブ系列に達するまで、ステップS713及びステップS714を実行する。その後、ステップS715に進む。

（ステップS713）重み算出部2061等は、楽譜特徴量と音響特徴量をマッチング処理する。重み算出部2061等は、例えば後述するマッチング処理を行って、両特徴量間における累積距離 $D_{n_2}$ を算出し、楽譜始点を示す対象楽譜サブ系列 $n_1$ 及び音響始点を示す対象音響サブ系列 $m_1$ を決定する。決定された楽譜特徴量と音響特徴量とのマッチング処理の詳細については後述する。その後、ステップS714に進む。

30

（ステップS714）重み算出部2061等は、楽譜終点 $n_2$ を直前のサブ系列 $n_2 - 1$ にシフトさせる。その後、ステップS712に進む。

（ステップS715）重み算出部2061等は、累積距離 $D_{n_2}$ が最小となる楽譜始点 $n_1$ から楽譜終点 $n_2$ までの楽譜特徴量を、音響特徴量と合致する区間と決定する。また、重み算出部1061等は、累積距離 $D_{n_2}$ が最小となる音響始点 $m_1$ から音響終点 $m_2$ までの音響特徴量を、楽譜特徴量と合致する音響特徴量の区間と決定する。その後、図4のステップS2に進む。

40

これにより、重み算出部2061は、楽譜特徴量と合致する音響特徴量の時系列の区間を観測時間として重み係数 $w_i$ を算出することになる。

【 0 1 2 6 】

次に、楽譜特徴量と音響特徴量とのマッチング処理について説明する。

図19は、本実施形態に係る楽譜特徴量と音響特徴量とのマッチング処理を示すフローチャートである。

（ステップS721）重み算出部2061等は、音響終点 $m_2$ を、例えば音響特徴量の時系列データにおける最後のフレームを含むサブ系列に初期設定する。重み算出部2061等は、対象楽譜サブ系列 $n_1$ を楽譜終点 $n_2$ と初期設定し、対象音響サブ系列 $m_1$ を音響終点 $m_2$ と初期設定する。ここで、対象楽譜サブ系列 $n_1$ とは、楽譜特徴量の時系列デー

50

タのうち処理対象となるサブ系列を示す。対象音響サブ系列  $m_1$  とは、音響特徴量の時系列データのうち処理対象となるサブ系列を示す。その後、ステップ S 7 2 2 に進む。

このように最後のサブ系列からマッチング処理を行うのは、楽譜位置推定装置 2 が最新の音響信号に基づいて楽譜位置を推定できるようにするためである。

【 0 1 2 7 】

(ステップ S 7 2 2) 重み算出部 2 0 6 1 等は、累積距離  $D_{n_2}$  の初期値を対象楽譜サブ系列  $n_1$  における楽譜特徴量と対象音響サブ系列  $m_1$  における音響特徴量との距離  $d_{n_1, m_1}$  と初期設定する。距離  $d_{n_1, m_1}$  は、楽譜特徴量と音響特徴量の類似度又は相関性を示す尺度である。距離  $d_{n_1, m_1}$  は、両特徴量の類似度又は相関性が高いほど小さい値をとる関数であればよい。例えば、距離  $d_{n_1, m_1}$  は、クロマベクトル重み係数  $w^{ch}$  に基づくエントロピー  $-w^{ch} \log w^{ch}$ 、周波数特性重み係数  $w_i^{sp}$  に基づくエントロピー  $-w^{sp} \log w^{sp}$ 、又はこれらの合計値でもよい。ここで、重み算出部 2 0 6 1 等は、例えば、クロマベクトル重み係数  $w^{ch}$  を算出する際、対象楽譜サブ系列  $n_1$  又は対象音響サブ系列  $m_1$  に相当する時間を観測時間として式 (4) 及び (5) を用いてもよい。また、重み算出部 1 0 6 1 等は、例えば、周波数特性重み係数  $w_i^{sp}$  を算出する際、対象楽譜サブ系列  $n_1$  又は対象音響サブ系列  $m_1$  に相当する時間を観測時間として式 (5) - (8) を用いてもよい。その後、ステップ S 7 2 3 に進む。

10

【 0 1 2 8 】

(ステップ S 7 2 3) 重み算出部 2 0 6 1 等は、対象楽譜サブ系列  $n_1$  が楽譜特徴量の時系列データにおける最初のサブ系列に達する、又は対象音響サブ系列  $m_1$  が音響特徴量の時系列データにおける最初のサブ系列に達するまで、ステップ S 7 2 4 ~ ステップ S 7 2 6 の処理を繰り返す。

20

(ステップ S 7 2 4) 重み算出部 2 0 6 1 等は、次の 3 通り (1) - (3) のサブ系列の組それぞれについて距離を算出する。重み算出部 2 0 6 1 等が、3 通りの距離を算出するのは、演奏される音楽におけるテンポのゆらぎが、楽譜特徴量を相対的に遅延させる場合及び音響特徴量を相対的に遅延させる場合を考慮するためである。

(1) 直前の対象楽譜サブ系列  $n_1 - 1$  における楽譜特徴量と現在の対象音響サブ系列  $m_1$  における音響特徴量との間の距離  $d_{n_1 - 1, m_1}$ 、

(2) 直前の対象楽譜サブ系列  $n_1 - 1$  における楽譜特徴量と直前の対象音響サブ系列  $m_1 - 1$  における音響特徴量との間の距離  $d_{n_1 - 1, m_1 - 1}$ 、及び

30

(3) 現在の対象楽譜サブ系列  $n_1$  における楽譜特徴量と直前の対象音響サブ系列  $m_1 - 1$  における音響特徴量との間の距離  $d_{n_1, m_1 - 1}$ 。

このステップにおいて、距離を算出する方法はステップ S 7 2 2 と同様な方法でよい。

重み算出部 1 0 6 1 等は、 $d_{n_1 - 1, m_1}$ 、 $d_{n_1 - 1, m_1 - 1}$ 、及び  $d_{n_1, m_1 - 1}$  のうち、最小の距離を選択する。その後、ステップ S 7 2 5 に進む。

【 0 1 2 9 】

(ステップ S 7 2 5) 重み算出部 2 0 6 1 等は、累積距離  $D_{n_2}$  に選択された距離を加算する。その後、ステップ S 7 2 6 に進む。

(ステップ S 7 2 6) 重み算出部 2 0 6 1 等は、対象楽譜サブ系列  $n_1$  及び対象音響サブ系列  $m_1$  を選択された距離に対応するサブ系列の組に変更する。つまり、距離  $d_{n_1 - 1, m_1}$  が選択された場合、重み算出部 2 0 6 1 等は、対象楽譜サブ系列  $n_1$  を直前の対象楽譜サブ系列  $n_1 - 1$  に変更し、対象音響サブ系列  $m_1$  を変更しない。

40

距離  $d_{n_1 - 1, m_1 - 1}$  が選択された場合、重み算出部 2 0 6 1 等は、対象楽譜サブ系列  $n_1$  を直前の対象楽譜サブ系列  $n_1 - 1$  に変更し、対象音響サブ系列  $m_1$  を直前の対象音響サブ系列  $m_1 - 1$  に変更する。

距離  $d_{n_1, m_1 - 1}$  が選択された場合、重み算出部 2 0 6 1 等は、対象楽譜サブ系列  $n_1$  を変更せず、対象音響サブ系列  $m_1$  を直前の対象音響サブ系列  $m_1 - 1$  に変更する。

その後、ステップ S 7 2 3 に進む。

【 0 1 3 0 】

50

これにより、重み算出部 2061 等は、音響特徴量と楽譜特徴量が合致する区間を特定して、重み算出部 2061 は、特定された区間に基づき重み係数  $w_i$  を算出する。よって、本実施形態に係る楽譜位置推定装置 2 は、さらに楽譜位置の推定誤りから容易に回復し、推定楽譜位置の出力を再開することができる。

#### 【0131】

なお、上述の観測時間を変更する処理において、観測時間を全区間に延長 (C.) した後 D T W 法を用いて楽譜特徴量と音響特徴量が合致する区間を特定することを前提にした説明したが、本実施形態ではこれには限られない。本実施形態では、シフト (A.) した後や、先行時間 T だけ延長 (B.) した後の観測時間における楽譜特徴量と音響特徴量の時系列データに D T W 法を用いて合致する区間を特定してもよい。これにより、逐次に楽譜位置の推定誤差が精度よく算出できる状態を回復することができる。

10

本実施形態では、観測時間を延長 (B.) する際の延長時間は先行時間 T に限られず、任意の時間でよいし、始点を固定した時刻にせず先行させてもよい。

なお、本実施形態では、より大きいサブ系列 (マッチング処理を行う区間数がより少ない) に基づき D T W 法を実行して楽譜特徴量を基準に合致する音響特徴量の区間の候補を決定しておき、その決定された区間の候補に対し、より小さいサブ系列 (マッチング処理を行う区間数がより多い) に基づき D T W 法を実行してもよい。

また、上述の D T W 法においては、楽譜特徴量を基準に合致する音響特徴量の区間を決定していたが、本実施形態では、音響特徴量を基準に合致する楽譜特徴量の区間を決定してもよい。

20

#### 【0132】

本実施形態において、周波数特性分析部 1021 は、入力されたデジタル音響信号が周波数領域信号に変換できれば、FFT 以外の方式、例えば DFT (Discrete Fourier Transform、離散フーリエ変換) を用いてもよい。

本実施形態において、再標本化部 1062 は、最も規格化粒子重み係数  $p_i$  が小さいものから予め定めた数  $N_r$  (例えば 50 個) の粒子を棄却してもよい。

#### 【0133】

なお、上述した実施形態における楽譜位置推定装置 1 又は 2 の一部、例えば、音響特徴量生成部 102、楽譜特徴量生成部 105、楽譜位置推定部 106、206、及び信頼度算出部 107、207 をコンピュータで実現するようにしても良い。その場合、この制御機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現しても良い。なお、ここでいう「コンピュータシステム」とは、楽譜位置推定装置 1 又は 2 に内蔵されたコンピュータシステムであって、OS や周辺機器等のハードウェアを含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM 等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでも良い。また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであっても良い。

30

40

また、上述した実施形態における楽譜位置推定装置 1 又は 2 の一部、または全部を、LSI (Large Scale Integration) 等の集積回路として実現しても良い。楽譜位置推定装置 1 又は 2 の各機能ブロックは個別にプロセッサ化してもよいし、一部、または全部を集積してプロセッサ化しても良い。また、集積回路化の手法は LSI に限らず専用回路、または汎用プロセッサで実現しても良い。また、半導体技術の進歩により LSI に代替する集積回路化の技術が出現した場合、当該技術による集積回路を用

50

いても良い。

【0134】

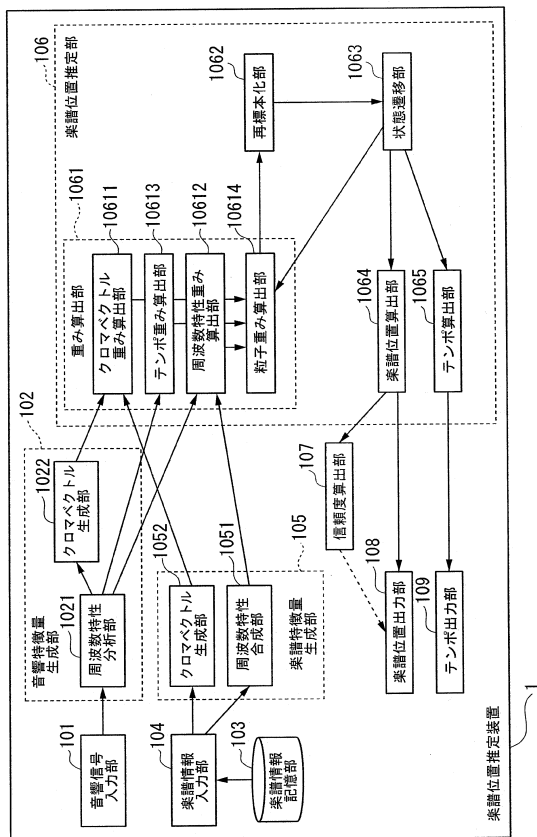
以上、図面を参照してこの発明の一実施形態について詳しく説明してきたが、具体的な構成は上述のものに限られることはなく、この発明の要旨を逸脱しない範囲内において様々な設計変更等を行うことが可能である。

【符号の説明】

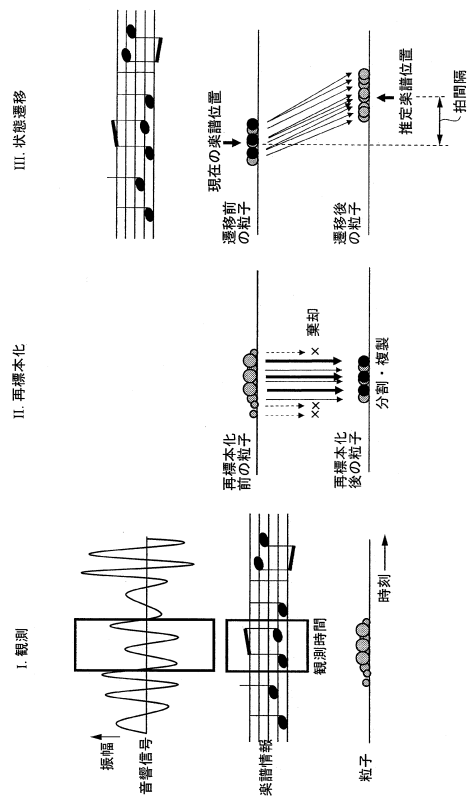
【0135】

- 1、2...楽譜位置推定装置、101...音響信号入力部、102...音響特徴量生成部、
- 1021...周波数特性分析部、1022...クロマベクトル生成部、
- 103...楽譜情報記憶部、104...楽譜情報入力部、
- 105...楽譜特徴量生成部、1051...周波数特性合成部、
- 1052...クロマベクトル生成部、
- 106、206...楽譜位置推定部、1061、2061...重み算出部、
- 10611...クロマベクトル重み算出部、10612...周波数特性重み算出部、
- 10613...テンポ重み算出部、10614...粒子重み算出部、
- 1062...再標本化部、1063...状態遷移部、1064...楽譜位置算出部、
- 1065...テンポ算出部、
- 107、207...信頼度算出部、108...楽譜位置出力部、109...テンポ出力部

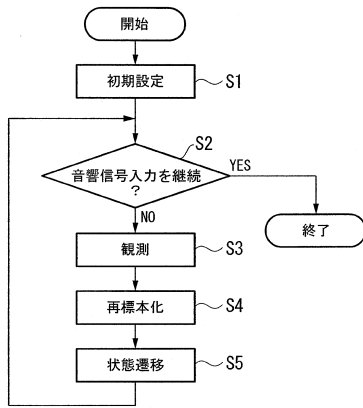
【図1】



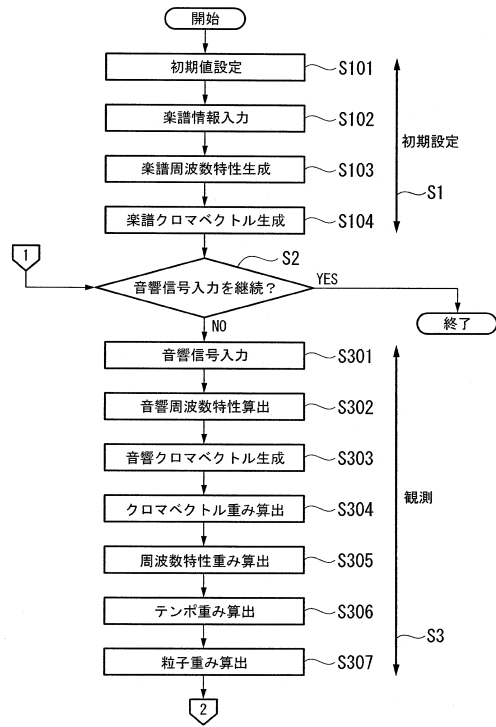
【図2】



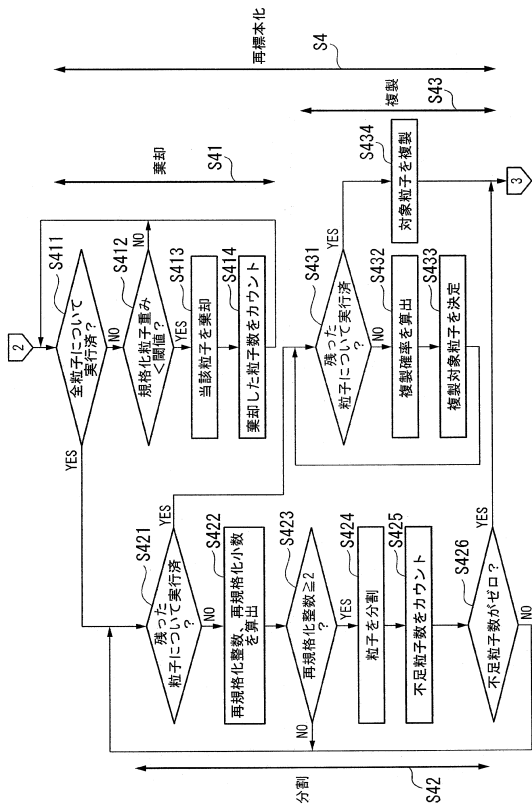
【図3】



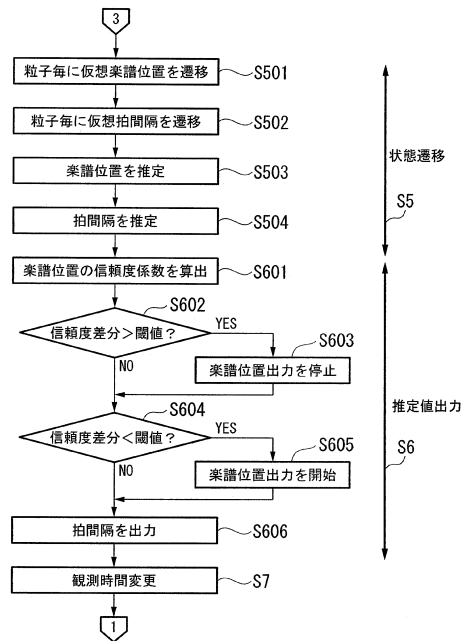
【図4】



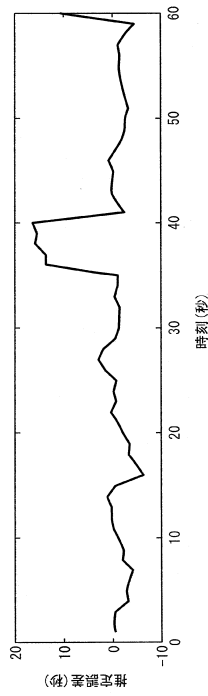
【図5】



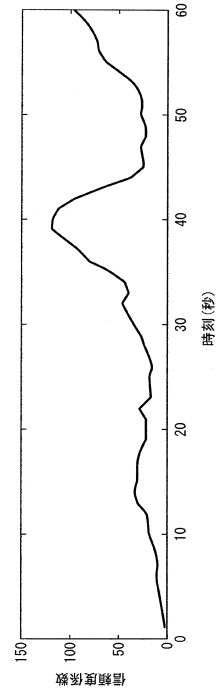
【図6】



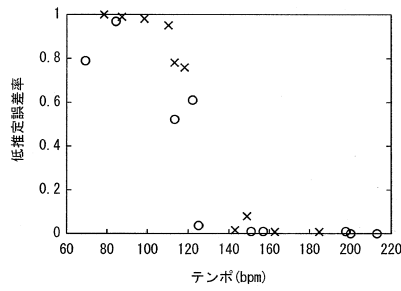
【図7】



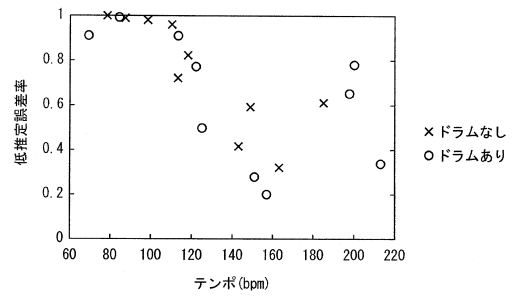
【図8】



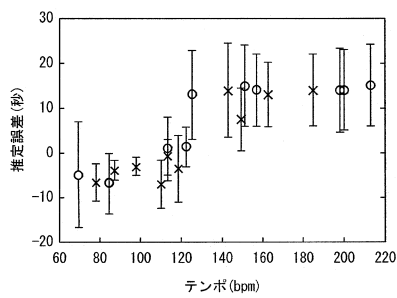
【図9】



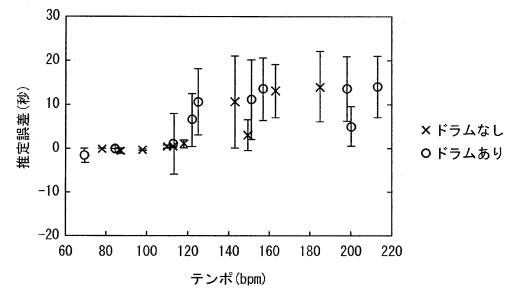
【図11】



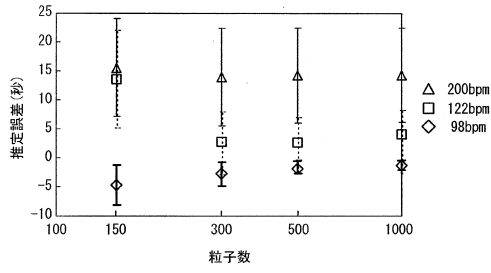
【図10】



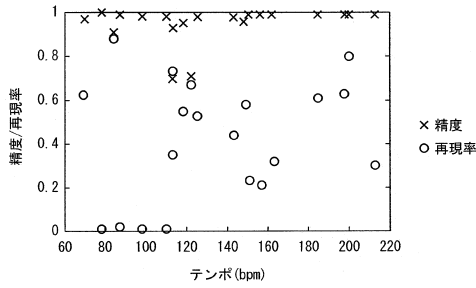
【図12】



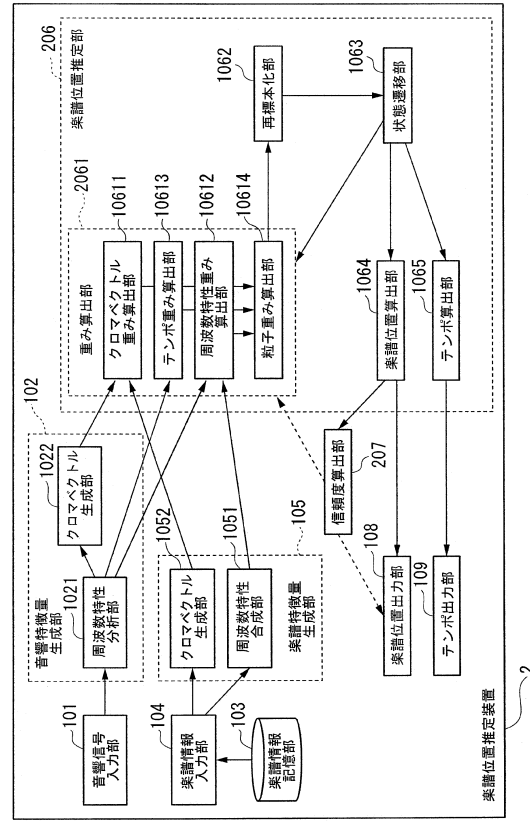
【図13】



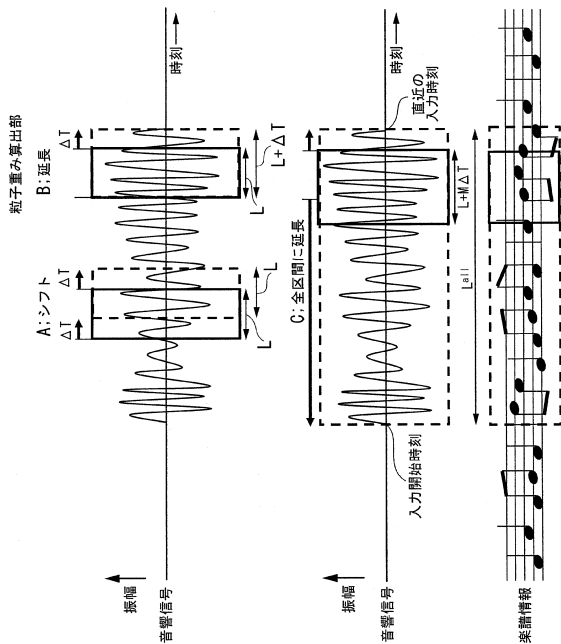
【図14】



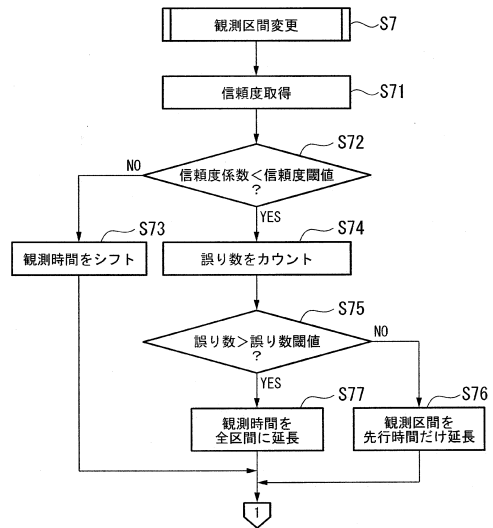
【図15】



【図16】

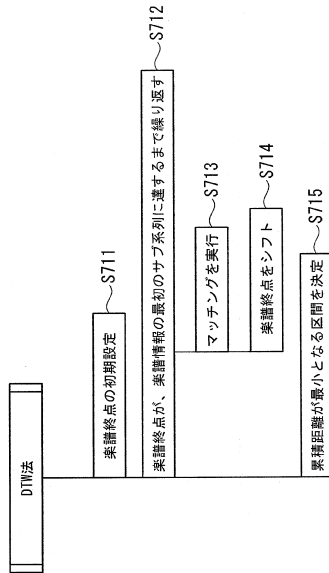


【図17】

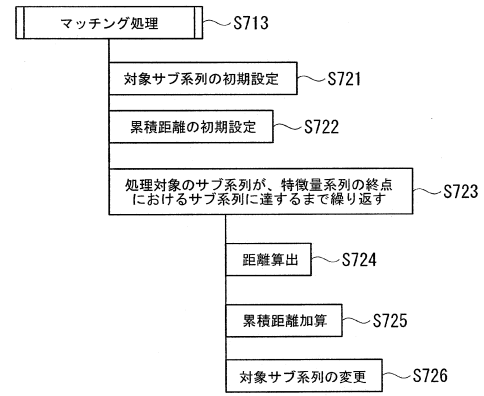




【図18】



【図19】



## フロントページの続き

- (72)発明者 中臺 一博  
埼玉県和光市本町8-1 株式会社ホンダ・リサーチ・インスティテュート・ジャパン内
- (72)発明者 大塚 琢馬  
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
- (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

審査官 山下 剛史

- (56)参考文献 特開2001-117580(JP,A)  
特開2007-241181(JP,A)  
特開2006-201278(JP,A)  
特開平5-27670(JP,A)  
米国特許第5913259(US,A)  
国際公開第02/47066(WO,A1)  
堀内靖雄他, "伴奏システム", 情報処理, 社団法人情報処理学会, 1994年 9月, Vol.35, No.9, pp.815-821  
浜中雅俊他, "発音時刻の楽譜上の位置を確率モデルにより推定するクオンタイズ手法", 情報処理学会論文誌, 2002年 2月, Vol.43, No.2, pp.234-244

(58)調査した分野(Int.Cl., DB名)

G10G 3/00 - 3/04

G10H 1/00