

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2017 年 4 月 6 日(06.04.2017)



(10) 国際公開番号
WO 2017/056885 A1

- (51) 国際特許分類:
G10L 25/54 (2013.01) *G10L 25/51* (2013.01)
G10H 1/00 (2006.01) *G10L 25/90* (2013.01)
- (21) 国際出願番号: PCT/JP2016/076266
- (22) 国際出願日: 2016 年 9 月 7 日(07.09.2016)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2015-194638 2015 年 9 月 30 日(30.09.2015) JP
- (71) 出願人: ヤマハ株式会社(YAMAHA CORPORATION) [JP/JP]; 〒4308650 静岡県浜松市中区中沢町 10 番 1 号 Shizuoka (JP).
- (72) 発明者: 高野 秀樹(TAKANO, Hideki); 〒4308650 静岡県浜松市中区中沢町 10 番 1 号 ヤマハ株式会社内 Shizuoka (JP).
- (74) 代理人: 大林 章, 外(OHBAYASHI, Akira et al.); 〒1130033 東京都文京区本郷 2-15-13 お茶の水ウイングビル 6 階 Tokyo (JP).

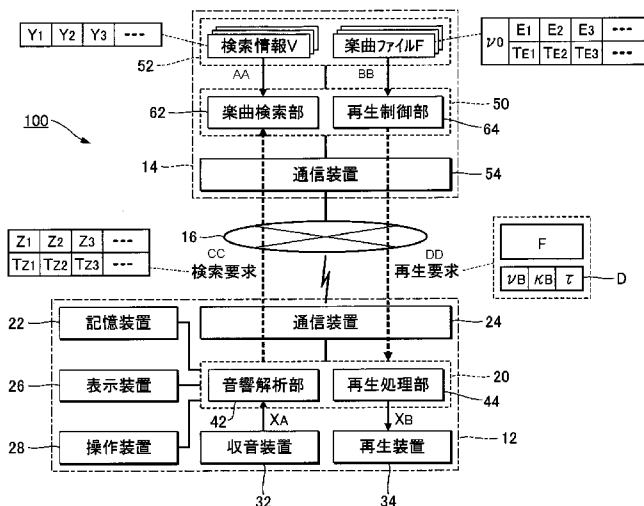
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

(54) Title: MUSIC PROCESSING METHOD AND MUSIC PROCESSING DEVICE

(54) 発明の名称: 楽曲処理方法および楽曲処理装置



22 Storage device
24 Communication device
26 Display device
28 Operation device
32 Recording device
34 Playback device
42 Acoustic analysis unit
44 Playback processing unit
54 Communication device
62 Music search unit
64 Playback control unit
AA Search information V
BB Music file F
CC Search request
DD Playback request

(57) Abstract: By music search processing that utilizes pitch information corresponding to the pitch of a target sound, this computer system searches for target music corresponding to the target sound while said target sound progresses, and, from the part of the target music corresponding to the progress of the target sound, causes the playback device to play back the target music while the target sound progresses.

(57) 要約: コンピュータシステムが、対象音の音高に応じた音高情報を利用した楽曲検索処理により、前記対象音に対応する目標楽曲を当該対象音の進行に並行して検索し、前記目標楽曲のうち当該対象音の進行に対応した部分から前記対象音の進行に並行して前記目標楽曲を再生装置に再生させる。

WO 2017/056885 A1

明 細 書

発明の名称： 楽曲処理方法および楽曲処理装置

技術分野

[0001] 本発明は、楽曲を検索および再生する技術に関する。

背景技術

[0002] 楽曲の歌唱により利用者が発音する音声（以下「歌唱音声」という）を解析することで、当該利用者が歌唱した楽曲を複数の候補から検索する技術が従来から提案されている。例えば特許文献1には、利用者の歌唱音声から順次に検出される基本周波数を中間的なフォーマットに変換してから複数の候補楽曲の各々との間で類似性を解析することで、歌唱音声に対応する楽曲を検索する技術が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2002-132278号公報

発明の概要

発明が解決しようとする課題

[0004] しかし、特許文献1の技術では、歌唱音声に対応する楽曲が検索されるに過ぎない。したがって、楽曲の再生に並行した歌唱を所望する利用者は、第1に、楽曲の一部を歌唱して当該楽曲を検索したうえで、第2に、当該楽曲を最初から再生して歌唱する必要がある。なお、以上の説明では楽曲の歌唱を便宜的に例示したが、例えば楽曲を楽器で演奏する場面でも同様の問題が発生し得る。以上の事情を考慮して、本発明は、楽曲の検索と当該楽曲の再生に並行した音響の発音（例えば当該楽曲の歌唱または演奏）とを所望する利用者の負担を軽減することを目的とする。

課題を解決するための手段

[0005] 本発明の好適な態様に係る楽曲処理方法は、コンピュータシステムが、音高が時間的に変化する対象音の音高に応じた音高情報を利用した楽曲検索処

理により、前記対象音に対応する目標楽曲を当該対象音の進行に並行して検索し、前記目標楽曲のうち当該対象音の進行に対応した部分から前記対象音の進行に並行して前記目標楽曲を再生装置に再生させる。

[0006] 本発明の好適な態様に係る楽曲処理装置は、音高が時間的に変化する対象音の音高に応じた音高情報を利用した楽曲検索処理により、対象音に対応する目標楽曲を当該対象音の進行に並行して検索する楽曲検索部と、目標楽曲のうち当該対象音の進行に対応した部分から対象音の進行に並行して目標楽曲を再生装置に再生させる再生制御部とを具備する。

図面の簡単な説明

[0007] [図1]本発明の第1実施形態に係る楽曲処理システムの構成図である。

[図2]音響解析処理のフローチャートである。

[図3]音響解析処理の説明図である。

[図4]楽曲検索処理のフローチャートである。

[図5]音符対応情報の説明図である。

[図6]再生制御処理のフローチャートである。

[図7]再生速度の設定の説明図である。

[図8]再生開始点の設定の説明図である。

[図9]楽曲処理システムの動作の説明図である。

[図10]変形例に係る楽曲処理装置の構成図である。

発明を実施するための形態

[0008] <第1実施形態>

図1は、本発明の第1実施形態に係る楽曲処理システム100の構成図である。図1に例示される通り、第1実施形態の楽曲処理システム100は、端末装置12と楽曲処理装置14とを具備する。端末装置12と楽曲処理装置14とは、移動体通信網またはインターネット等を含む通信網16を介して相互に通信する。端末装置12は、例えば携帯電話機もしくはスマートフォン等の可搬型の通信端末、またはパーソナルコンピュータ等の可搬型もしくは据置型の通信端末である。楽曲処理装置14は、端末装置12の利用者

が歌唱した楽曲（以下「目標楽曲」という）を検索するとともに目標楽曲を端末装置１２に再生させるサーバ装置である。第１実施形態の楽曲処理装置１４は、端末装置１２から送信される検索要求（クエリ）を契機として目標楽曲を検索し、当該目標楽曲の楽曲ファイルＦを含む再生要求を送信することで端末装置１２に目標楽曲を再生させる。楽曲ファイルＦは、例えば目標楽曲の伴奏音を表すデータである。なお、実際には複数の端末装置１２が楽曲処理装置１４と通信し得るが、以下の説明では便宜的に１個の端末装置１２に着目する。

[0009] 図１に例示される通り、端末装置１２は、制御装置２０と記憶装置２２と通信装置２４と表示装置２６と操作装置２８と收音装置３２と再生装置３４とを具備するコンピュータシステムで実現される。制御装置２０は、例えばＣＰＵ（Central Processing Unit）で構成され、端末装置１２の各要素を統括的に制御する。記憶装置２２は、例えば磁気記録媒体および半導体記録媒体等の公知の記録媒体、または、複数種の記録媒体の組合せで実現され、制御装置２０が実行するプログラムと制御装置２０が使用する各種のデータとを記憶する。通信装置２４は、通信網１６を介して楽曲処理装置１４と通信する。第１実施形態の通信装置２４は、目標楽曲の検索要求を楽曲処理装置１４に送信し、楽曲処理装置１４から送信された目標楽曲の楽曲ファイルＦを受信する。

[0010] 表示装置２６は、制御装置２０から指示された画像を表示する。例えば楽曲処理装置１４による検索結果（目標楽曲の楽曲名）が表示装置２６に表示される。操作装置２８は、利用者からの指示を受付ける入力機器である。操作装置２８は、例えば利用者による操作を検知する複数の操作子、または、表示装置２６の表示面に対する利用者の接触を検知するタッチパネルである。

[0011] 收音装置３２は、周囲の音響を收音して音響信号XAを生成する。端末装置１２の利用者は、收音装置３２に向かって所望の楽曲を歌唱する。收音装置３２は、端末装置１２の利用者が楽曲を歌唱した音声（以下「歌唱音声」と

いう)を表す音響信号XAを生成する。なお、音響信号XAをアナログからデジタルに変換するA/D変換器の図示は便宜的に省略した。

[0012] 図1に例示される通り、第1実施形態の制御装置20は、記憶装置22に記憶されたプログラムを実行することで音響解析部42および再生処理部44として機能する。音響解析部42は、收音装置32から供給される音響信号XAを解析することで、歌唱音声に含まれる音符毎の音高に応じた音高情報Zを生成する。再生処理部44は、通信装置24が楽曲処理装置14から受信した楽曲ファイルFから目標楽曲の音響信号XBを生成して再生装置34に供給する。

[0013] 再生装置34(例えばスピーカまたはヘッドホン)は、再生処理部44から供給される音響信号XBが示す楽曲を再生する。具体的には、目標楽曲の伴奏音(例えば楽器の演奏音)が再生装置34により再生される。なお、音響信号XBをデジタルからアナログに変換するD/A変換器の図示は便宜的に省略した。

[0014] 図2は、第1実施形態の音響解析部42が音高情報Zを生成する処理(以下「音響解析処理」という)のフローチャートである。端末装置12に対する利用者からの指示を契機として音響解析処理が開始される。音響解析処理を開始すると、音響解析部42は、音響信号XAの音高(基本周波数)pを順次に検出し(SA1)、図3に例示される通り、音響信号XAを時間軸上で安定区間x1と変動区間x2とに区分する(SA2)。安定区間x1は、音高pが時間的に安定に推移する区間であり、変動区間x2は、音高pが不安定に変動する区間である。具体的には、複数の音高pに関する特徴量の時間変化が特定の条件を満たす一連の区間(例えば複数の音高pの分散等の散布度が所定値を下回る一連の区間)が安定区間x1として画定される。音響信号XAの任意の1個の安定区間x1は、楽曲の1個の音符に対応する。音響解析部42は、安定区間x1内の複数の音高pの代表値(例えば平均値または中央値)を当該安定区間x1の音高Pとして安定区間x1毎(すなわち楽曲の音符毎)に算定する(SA3)。

[0015] 音響解析部42は、音響信号XAの音符毎の音高Pに応じた音高情報Zを順次に生成する(SA4)。図3に例示される通り、第1実施形態の音響解析部42は、音響信号XAの1個の安定区間x1と直前の安定区間x1との間における音高Pの差分値(すなわち音高差) δ に応じた音高情報Zを安定区間x1毎に生成する。具体的には、音高差 δ の数値範囲を所定の単位(例えば100cent単位)で区画した複数の範囲の各々に相異なる符号(例えば“A”~“Z”の文字)が付与され、複数の範囲のうち音高差 δ を内包する範囲に付与された符号が音高情報Zとして生成される。以上の説明から理解される通り、音高情報Zは歌唱音声の音高Pを相対値(直前の音高Pからの変動量)で表現する情報であり、音高Pの絶対値と安定区間x1の時刻および時間長(すなわち利用者による歌唱のリズム)とは音高情報Zに反映されない。

[0016] 音響解析部42は、音高情報Zの生成毎(すなわち安定区間x1毎)に検索要求を生成して通信装置24から楽曲処理装置14に送信する。検索要求は、図1に例示される通り、現時点までに生成された複数の音高情報Zの時系列(Z1, Z2, ……)と、各音高情報Zに対応する音符の時間軸上の位置を指定する時間情報TZ(TZ1, TZ2, ……)とを包含する。任意の1個の音符の時間情報TZは、例えば当該音符の発音点の時刻(例えば安定区間x1の始点の時刻)を指定する。以上の説明から理解される通り、音高情報Zと時間情報TZとを音符毎に含む検索要求が、歌唱音声の進行に並行して端末装置12から楽曲処理装置14に順次に送信される。

[0017] なお、音響信号XAには、音高pが適正に検出されない区間(以下「非検出区間」という)が含まれ得る。例えば、音量の不足等の理由により音高pを正確に検出できない無音区間と、調波構造を持たない子音が発音されている子音区間とが、非検出区間として想定される。非検出区間の直前の区間aと直後の区間bとで音高pが同一である場合、区間aとその直前の区間との音高差に応じた音高情報Zと、区間bとその直前に音高が検出された区間aとの音高差(すなわちゼロ)に応じた音高情報Zとを含む検索要求が生成される。また、無音区間または子音区間について、非検出区間であることを示す

記号を音高情報Zとして生成することも可能である。また、1個の子音に対応する子音区間を、当該子音の直後の母音の区間に含めて1個の音符として、時間情報TZを決定することも可能である。

[0018] 図1に例示される通り、楽曲処理装置14は、制御装置50と記憶装置52と通信装置54とを具備するコンピュータシステムで実現される。制御装置50は、例えばCPUで構成され、楽曲処理装置14の各要素を統括的に制御する。通信装置54は、通信網16を介して端末装置12と通信する。第1実施形態の通信装置54は、端末装置12から送信された検索要求を受信するとともに、当該検索要求に応じて検索された目標楽曲の楽曲ファイルFを要求元の端末装置12に送信する。記憶装置52は、例えば磁気記録媒体および半導体記録媒体等の公知の記録媒体、または、複数種の記録媒体の組合せで実現され、制御装置50が実行するプログラムと制御装置50が使用する各種のデータとを記憶する。

[0019] 図1に例示される通り、第1実施形態の記憶装置52は、目標楽曲の候補となるN個の楽曲（以下「候補楽曲」という）の各々について検索情報Vと楽曲ファイルFとを記憶する（Nは2以上の自然数）。検索情報Vは、目標楽曲の検索に利用される情報であり、候補楽曲内の音符毎の音高の時系列を表現する。図1に例示される通り、各候補楽曲の検索情報Vは、当該候補楽曲の相異なる音符に対応する複数の音高情報Yの時系列（Y1, Y2, ……）で構成される。任意の1個の音符の音高情報Yは、端末装置12の音響解析部42が生成する前述の音高情報Zと同様に、直前の音符との音高差に応じた符号（例えば“A”～“Z”の文字）を指定する。

[0020] 他方、各候補楽曲の楽曲ファイルFは、当該候補楽曲の演奏内容を指定する。具体的には、第1実施形態の楽曲ファイルFは、候補楽曲の伴奏音を構成する各音符の発音または消音を指示する指示情報E（E1, E2, ……）と、各指示情報Eの処理時点（例えば相前後する指示の時間間隔）を指定する時間情報TE（TE1, TE2, ……）とが時系列に配列されたMIDI（Musical Instrument Digital Interface）形式のファイルである。ただし、MIDI

l 形式は楽曲ファイルFの形式の好適な例示であり、楽曲ファイルFの形式は任意である。また、楽曲ファイルFは、候補楽曲の標準的な再生速度（以下「標準再生速度」という） ν_0 を包含する。

[0021] 図1に例示される通り、第1実施形態の制御装置50は、記憶装置52に記憶されたプログラムを実行することで楽曲検索部62および再生制御部64として機能する。楽曲検索部62は、端末装置12の利用者が発音した歌唱音声に対応する目標楽曲を検索する。再生制御部64は、楽曲検索部62が検索した目標楽曲を端末装置12の再生装置34に再生させる。具体的には、再生制御部64は、目標楽曲の楽曲ファイルFを含む再生要求を検索要求の送信元の端末装置12に通信装置54から送信する。

[0022] 図4は、第1実施形態の楽曲検索部62が目標楽曲を検索する処理（以下「楽曲検索処理」という）のフローチャートである。通信装置54による検索要求の受信を契機として楽曲検索処理が開始される。

[0023] 楽曲検索処理を開始すると、楽曲検索部62は、N個の候補楽曲から1個の候補楽曲（以下「選択候補楽曲」という）を選択する（SB1）。そして、楽曲検索部62は、歌唱音声と選択候補楽曲とを相互に対比する第1処理SB2により歌唱音声と選択候補楽曲との類否の指標（以下「類似指標」という）Mを算定する。第1処理SB2は、選択候補楽曲の検索情報Vが指定する音高情報Yの時系列（符号列）と、検索要求で指定される歌唱音声の音高情報Zの時系列とを相互に対比する処理である。

[0024] 第1実施形態の類似指標Mは、選択候補楽曲の音高情報Yの時系列と歌唱音声の音高情報Zの時系列との間の編集距離 ε に応じて設定される。編集距離 ε は、音高情報Yまたは音高情報Zの編集（削除、挿入、置換）により音高情報Yの時系列および音高情報Zの時系列の一方を他方に変換するときの編集回数の最小値を意味する。したがって、歌唱音声と選択候補楽曲とが類似するほど類似指標Mは小さい数値となる。具体的には、楽曲検索部62は、選択候補楽曲を時間軸上で区画した複数の区間の各々について当該区間内の音高情報Yの時系列と歌唱音声の音高情報Zの時系列との編集距離 ε を算

定し、編集距離 ε の最小値（すなわち選択候補楽曲のうち歌唱音声に最も類似する区間の編集距離 ε ）を当該選択候補楽曲の類似指標 M として確定する。編集距離 ε の算定には、例えば特開 2013-242409 号公報にも開示される通り、例えば動的計画法（Dynamic Programming）が好適に利用される。

[0025] 第 1 実施形態の第 1 処理 SB2 は、選択候補楽曲の各音高情報 Y と歌唱音声の各音高情報 Z との時間的な対応（すなわち、選択候補楽曲と歌唱音声とにおける各音符の時間的な対応）を示す経路を探索する処理（すなわち経路探索処理）である。経路探索処理のなかで類似指標 M も算定される。すなわち、選択候補楽曲と歌唱音声との間の各音符の時間的な対応を表す音符対応情報 C が類似指標 M の算定とともに生成される。具体的には、図 5 に例示される通り、楽曲検索部 62 は、選択候補楽曲の各音高情報 Y と歌唱音声の各音高情報 Z とを相互に対応させる最適経路を動的計画法による編集距離 ε の算定とともに探索し、当該最適経路上の音高情報 Y と音高情報 Z との組毎に選択候補楽曲の各音符と歌唱音声の各音符とを対応させた音符対応情報 C を生成する。なお、音符対応情報 C の解析には、ビタビアルゴリズム等の公知の経路探索処理も採用され得る。

[0026] 楽曲検索部 62 は、 N 個の候補楽曲について類似指標 M および音符対応情報 C の生成（すなわち第 1 処理 SB2）を実行したか否かを判定する（SB3）。判定結果が否定である場合（SB3 : NO）、楽曲検索部 62 は、 N 個の候補楽曲のうち未処理の候補楽曲を新たな選択候補楽曲として選択（SB1）したうえで当該選択候補楽曲について第 1 処理 SB2 を実行する。他方、 N 個の候補楽曲について第 1 処理 SB2 を実行した場合（SB3 : YES）、楽曲検索部 62 は、各候補楽曲の類似指標 M に応じて N 個の候補楽曲から目標楽曲を選択する第 2 処理 SC を実行する。

[0027] 具体的には、楽曲検索部 62 は、 N 個の候補楽曲のうち類似指標 M が最小値（ $M_{\min}$ ）である 1 個の候補楽曲を暫定的な目標楽曲（以下「暫定楽曲」という）として特定し（SC1）、暫定楽曲の類似指標 $M_{\min}$ が閾値 M_{th} を下回る

か否かを判定する（SC2）。類似指標Mminが閾値Mthを下回る場合（SC2：YES）には、暫定楽曲が歌唱音声に十分に類似すると評価できるから、楽曲検索部62は、当該暫定楽曲を目標楽曲として確定する（SC3）。他方、類似指標Mminが閾値Mthを上回る場合（SC2：NO）には、暫定楽曲と歌唱音声とは必ずしも類似しないと評価できるから、目標楽曲を確定せずに楽曲検索処理を終了する。以上に例示した楽曲検索処理により、歌唱音声に類似する目標楽曲が検索されるとともに、歌唱音声と目標楽曲との間の各音符の時間的な対応を示す音符対応情報Cが生成される。

[0028] 図1の再生制御部64は、記憶装置52に記憶されたN個の楽曲ファイルFのうち楽曲検索部62が検索した目標楽曲の楽曲ファイルFを含む再生要求を端末装置12に送信することで、端末装置12の再生装置34に目標楽曲を再生させる。図6は、再生制御部64が目標楽曲を再生装置34に再生させるための処理（以下「再生制御処理」という）のフローチャートである。楽曲検索部62による目標楽曲の検索を契機として再生制御処理が開始される。

[0029] 再生制御処理を開始すると、再生制御部64は、楽曲検索部62が検索した目標楽曲の楽曲ファイルFを記憶装置52から取得し（SD1）、目標楽曲の再生条件を指定する再生情報Dを生成する（SD2～SD4）。再生情報Dは、楽曲ファイルFによる目標楽曲の再生を歌唱音声に追従させる（すなわち目標楽曲の再生が歌唱音声に同期するように目標楽曲を歌唱音声に応じた速度で再生する）ための再生条件を指定する情報であり、歌唱音声の特性に応じて可変に設定される。具体的には、第1実施形態の再生情報Dは、再生速度 ν_B と再生調 κ_B と時間対応情報 τ とを包含する。楽曲検索部62は、再生速度 ν_B の設定（SD2）と再生調 κ_B の設定（SD3）と時間対応情報 τ の設定（SD4）とを順次に実行する。そして、再生制御部64は、目標楽曲の楽曲ファイルFと再生情報Dとを含む再生要求を通信装置54から端末装置12に送信する（SD5）。再生情報Dの具体的な内容と設定方法について以下に例示する。

[0030] 再生速度 ν_B は、再生装置34による目標楽曲の再生の速度（テンポ）であり、例えば単位時間毎の拍数（Beats Per Minute）で表現される。図7は、再生制御部64が再生速度 ν_B を設定する処理（SD2）の説明図である。第1実施形態の再生制御部64は、歌唱音声の進行の速度（以下「発音速度」という） ν_A に近似または合致するように、発音速度 ν_A に応じて目標楽曲の再生速度 ν_B を設定する。

[0031] 図7には、歌唱音声の各音符の発音点（例えば安定区間 $\times 1$ の始点）と目標楽曲の楽曲ファイルFが指定する各音符（歌唱音声の各音符に対応する区間内）の発音点とが共通の時間軸のもとで図示されている。歌唱音声の各音符（音高情報Z）の発音点は検索要求内の時間情報TZ（TZ1, TZ2, ……）で指定され、目標楽曲の各音符（指示情報E）の発音点は楽曲ファイルF内の時間情報TE（TE1, TE2, ……）で指定される。歌唱音声の各音符と目標楽曲の各音符との時間的な対応は、前述の第1処理SB2により生成された音符対応情報Cで指定される。図7に例示される通り、楽曲ファイルFで指定される標準再生速度 ν_0 で目標楽曲を再生した場合、歌唱音声と目標楽曲との間では、音符対応情報Cが対応を指定する各音符の発音点の時間軸上の位置が相違し得る。第1実施形態の再生制御部64は、歌唱音声と目標楽曲との間で音符対応情報Cが指定する音符毎の時間的な対応を参照して、歌唱音声の発音速度 ν_A に対応する目標楽曲の再生速度 ν_B を設定する。

[0032] 再生制御部64は、まず、検索要求で指定された最後の音符（すなわち、歌唱者が歌唱した直前の音符）までの各音符について直前の音符との時間間隔（IOI：Inter Onset Interval）QA（QA1, QA2, ……）を算定し、当該最後の音符までの時間間隔QAの合計値LAを算定する（SD21）。また、再生制御部64は、目標楽曲のうち歌唱音声における歌唱済の区間に対応する各音符について直前の音符との時間間隔QB（QB1, QB2, ……）を算定し、複数の音符について時間間隔QBの合計値LBを算定する（SD22）。そして、再生制御部64は、楽曲ファイルFで指定される目標楽曲の標準再生速度 ν_0 に、合計値LAに対する合計値LBの比率（LB/LA）を乗算することで、目標

楽曲の再生速度 ν_B ($\nu_B = \nu_0 \times (L_B/L_A)$) を算定する (SD23)。すなわち、目標楽曲のうち歌唱音声の各音符に対応する区間が、時間軸上で比率 (L_B/L_A) に応じて伸縮される。したがって、歌唱音声の発音速度 ν_A が目標楽曲の標準再生速度 ν_0 と比較して速い ($L_A < L_B$) ほど、目標楽曲の再生速度 ν_B は増加する。なお、図 7 に例示される通り、歌唱音声の各音符の時間間隔 Q_{An} ($n = 1, 2, \dots$) と目標楽曲の各音符の時間間隔 Q_{Bn} とで相互に対応するもの同士の差分 Δn ($\Delta n = Q_{An} - Q_{Bn}$) の合計値 ($\Delta 1 + \Delta 2 + \dots$) が減少する (理想的にはゼロになる) ように、再生制御部 64 が目標楽曲の再生速度 ν_B を調整する、と表現することも可能である。

[0033] 他方、再生制御部 64 がステップ SD3 で設定する再生調 κ_B は、再生装置 34 による目標楽曲の再生の調 (キー) である。第 1 実施形態の端末装置 12 が送信する検索要求は、歌唱音声の最初の音符の音高 P (絶対値) を包含する。したがって、各音高情報 Z で指定される音高差を最初の音符の音高 P に順次に加算することで歌唱音声の音符毎の音高を特定することが可能である。なお、音符毎に音高情報 Z と音高 P とを検索要求に含めて端末装置 12 から楽曲処理装置 14 に送信することも可能である。再生制御部 64 は、歌唱音声と楽曲ファイル F が示す目標楽曲との間で音符対応情報 C が対応を示す音符間の音高差に応じて目標楽曲の再生調 κ_B を設定する。具体的には、歌唱音声の各音符の音高と楽曲ファイル F の各指示情報 E が示す目標楽曲の各音符の音高との音高差を複数の音符にわたり平均した数値が所定の閾値を下回る (例えばゼロとなる) ように、再生制御部 64 は再生調 κ_B を設定する。

[0034] また、再生制御部 64 は、図 6 のステップ SD4 において時間対応情報 τ を生成する。時間対応情報 τ は、歌唱音声の任意の 1 個の音符 (以下「基準音符」という) について時間情報 T_Z が示す時点と、目標楽曲のうち音符対応情報 C により基準音符に対応づけられた音符の時間情報 T_E が示す時点とを相互に対応させる情報である。すなわち、時間対応情報 τ は、歌唱音声における 1 個の時刻 (時間情報 T_Z が示す時点) と目標楽曲における 1 個の時刻 (時間情報 T_E が示す時点) との対応を示す情報である。時間情報 T_E が示す時点は

、例えば目標楽曲の始点からのM I D Iクロックの個数で表現される。

[0035] 以上に例示した再生速度 νB と再生調 κB と時間対応情報 τ とを含む再生情報Dが目標楽曲の楽曲ファイルFとともに再生要求として楽曲処理装置14から端末装置12に送信される。端末装置12の通信装置24が再生要求を受信すると、再生処理部44は、楽曲ファイルFで指定される音響の音響信号XBを再生情報Dに応じて生成して再生装置34に供給する。したがって、再生情報Dが指定する再生速度 νB および再生調 κB のもとで目標楽曲が再生装置34により再生される。すなわち、目標楽曲の楽曲ファイルFの再生の速度が、再生情報Dに含まれる再生速度 νB に調整され、楽曲ファイルFの再生の調が、再生情報Dに含まれる再生調 κB に調整される。

[0036] 端末装置12から検索要求が送信されてから現時点までの期間では利用者による歌唱が進行している。したがって、目標楽曲のうち検索要求の送信の時点で利用者が歌唱した区間の直後から目標楽曲の再生を開始したのでは、目標楽曲の再生時点が利用者による歌唱に対して遅延した状態となる。そこで、第1実施形態の再生処理部44は、歌唱音声の進行に対応した時点（すなわち、端末装置12の利用者が実際に歌唱している時点）から目標楽曲の再生が開始されるように、目標楽曲内の再生の開始の時点（以下「再生開始点」という）を時間対応情報 τ に応じて制御する。具体的には、図8に例示される通り、時間対応情報 τ が指定する基準音符の発音の時点（時間情報TZが示す時刻）から現在までの経過時間 σ にわたり、目標楽曲内で基準音符に対応する音符の時点 t_1 から再生速度 νB で再生点を時間的に進行させた時点 t_2 が、目標楽曲の再生開始点として選定される。したがって、利用者による歌唱音声と同等の再生速度 νB および再生調 κB のもとで、目標楽曲のうち利用者が歌唱した時点以降の区間が再生される。すなわち、歌唱音声に追従する（すなわち、歌唱音声に応じた速度で目標楽曲を再生することで目標楽曲が歌唱音声に同期する）ように目標楽曲が再生される。以上の説明から理解される通り、第1実施形態における楽曲処理装置14の再生制御部64は、歌唱音声の進行に対応した再生開始点 t_2 から目標楽曲を歌唱音声の進行に並

行して再生装置 34 に再生させる要素として機能する。

[0037] 図 9 は、楽曲処理システム 100 の動作の説明図である。図 9 に例示される通り、音響信号 XA の安定区間 $\times 1$ の検出毎（すなわち音符毎）に、複数の音高情報 Z の時系列と各音高情報 Z の時間情報 TZ とを包含する検索要求が、端末装置 12 から楽曲処理装置 14 に送信される（S1）。通信装置 54 が検索要求を受信すると、楽曲検索部 62 は、歌唱音声に対応する 1 個の目標楽曲を図 4 の楽曲検索処理により検索する（S2）。楽曲検索部 62 は、楽曲検索処理の結果の通知（以下「結果通知」という）を端末装置 12 に送信する（S3）。結果通知は目標楽曲の楽曲名を包含する。端末装置 12 の制御装置 20 が結果通知を受信すると、制御装置 20 は、結果通知で指定された目標楽曲の楽曲名を表示装置 26 に表示させて検索結果を利用者に提示する。なお、例えば目標楽曲の再生を予告する「約 3 秒後に楽曲 A を再生します」等のメッセージを目標楽曲の楽曲名とともに表示装置 26 に表示させることも可能である。

[0038] 利用者は、表示装置 26 の表示を確認し、目標楽曲が自身の歌唱中の楽曲であるか否か（検索結果の正誤）を判定する。そして、目標楽曲が歌唱中の楽曲でない場合（検索結果が誤りである場合）、利用者は、操作装置 28 を適宜に操作することで楽曲変更を指示する。図 9 に破線で図示される通り、楽曲変更の指示は、端末装置 12 から楽曲処理装置 14 に送信される。目標楽曲が歌唱中の楽曲である場合（検索結果が適正である場合）には、利用者は操作装置 28 を操作しない。したがって、楽曲処理装置 14 に楽曲変更の指示は送信されない。

[0039] 楽曲処理装置 14 の楽曲検索部 62 は、結果通知（S3）から所定時間（以下「待機時間」という）が経過するまでに利用者から楽曲変更が指示されたか否かを判定する（S4）。待機時間は、例えば数秒程度の時間長に設定される。待機時間内に楽曲変更が指示された場合（S4: YES）、楽曲検索部 62 は、直前の楽曲検索処理の結果を破棄し、端末装置 12 から以降に送信される検索要求（S1）を利用して楽曲検索処理を再実行する（S2）。検索要求

には送信毎に新たな音符の音高情報 Z が追加されるから、楽曲検索処理で検索される目標楽曲は楽曲検索処理毎に変化する。以上の説明から理解される通り、第1実施形態の楽曲検索部62は、歌唱音声の進行に並行して楽曲検索処理により目標楽曲を検索する。

[0040] 他方、楽曲変更を利用者から指示されずに待機時間が経過した場合には（ $S4:N0$ ）、直前の楽曲検索処理で検索された目標楽曲が検索結果として確定する。再生制御部64は、直前の楽曲検索処理で検索された目標楽曲を端末装置12の再生装置34に再生させる（ $S5$ ）。具体的には、再生制御部64は、目標楽曲の楽曲ファイル F と当該目標楽曲の再生条件（再生速度 νB 、再生調 κB 、時間対応情報 τ ）を指定する再生情報 D とを含む再生要求を、通信装置54から通信網16を介して端末装置12に送信する。端末装置12の再生処理部44は、目標楽曲のうち時間対応情報 τ と再生速度 νB とに応じた時点から再生速度 νB および再生調 κB で目標楽曲が再生されるように楽曲ファイル F から音響信号 XB を生成して再生装置34に供給する。すなわち、歌唱音声の進行に並行して歌唱音声に追従するように目標楽曲が再生される。

[0041] 以上に説明した通り、第1実施形態では、歌唱音声に対応する目標楽曲が歌唱音声の進行に並行して検索され、かつ、目標楽曲のうち歌唱音声の進行に対応した再生開始点 t_2 から歌唱音声の進行に並行して目標楽曲が再生される。すなわち、利用者が所望の楽曲を歌唱した一連の歌唱音声の進行に並行して目標楽曲の検索と当該目標楽曲の再生とが順次に実行される。具体的には、利用者が所望の楽曲を継続的に歌唱していると、途中の時点から歌唱に並行して目標楽曲（伴奏音）の再生が開始される。したがって、目標楽曲の検索と目標楽曲の再生に並行した当該目標楽曲の歌唱とを所望する利用者の負担を軽減することが可能である。

[0042] 第1実施形態では、楽曲検索処理の結果通知から所定の待機時間が経過するまでに利用者が楽曲変更を指示した場合に楽曲検索処理が再実行される。したがって、楽曲検索処理による誤検索を修正して適正な目標楽曲を再生することが可能である。また、楽曲変更が利用者から指示されずに待機時間が

経過した場合には、直前の楽曲検索処理で検索された目標楽曲の再生が開始されるから、利用者による歌唱を阻害せずに適正な目標楽曲を再生できるという利点がある。

[0043] 第1実施形態では、歌唱音声の発音速度 νA に対応する再生速度 νB で目標楽曲が再生されるから、楽曲検索処理の実行後に、目標楽曲の再生に並行した歌唱を利用者が違和感なく継続できるという利点がある。また、楽曲検索処理のうち第1処理SB2の結果（歌唱音声と目標楽曲との間の各音符の時間的な対応）が、目標楽曲の再生速度 νB の設定に流用される。したがって、目標楽曲の再生速度 νB を楽曲検索処理とは別個の処理で設定する構成と比較して、再生制御部64の処理負荷を軽減することが可能である。

[0044] 第1実施形態では、歌唱音声と目標楽曲との間で相互に対応する音符間の音高差に応じた再生調 κB で目標楽曲が再生されるから、楽曲検索処理の実行後に、目標楽曲の再生に並行した歌唱を利用者が違和感なく継続できるという利点がある。また、楽曲検索処理のうち第1処理SB2の結果（歌唱音声と目標楽曲との間の各音符の時間的な対応）が、目標楽曲の再生調 κB の設定に流用される。したがって、目標楽曲の再生調 κB を楽曲検索処理とは別個の処理で設定する構成と比較して、再生制御部64の処理負荷を軽減することが可能である。

[0045] ところで、目標楽曲の再生時点を端末装置12に指示する構成としては、例えば、音響処理装置14において再生開始点 t_2 を特定し、時間対応情報 τ の代わりに再生開始点 t_2 を含む再生情報Dを端末装置12に送信する構成（以下「対比例」という）も想定される。しかし、通信網16における伝送遅延は刻々と変動する。したがって、対比例では、伝送遅延の大小に起因して、例えば端末装置12の利用者が実際に歌唱している箇所に対して乖離した時点から目標楽曲の再生が開始される可能性がある。対比例とは対照的に、第1実施形態では、歌唱音声と目標楽曲との間の時間的な対応を指定する時間対応情報 τ と目標楽曲の再生速度 νB とが楽曲ファイルFとともに端末装置12に送信される。したがって、通信網16での伝送遅延の大小に関わらず

、目標楽曲のうち歌唱音声の進行に応じた再生開始点 t_2 を特定できる（歌唱音声に同期するように楽曲ファイル F の目標楽曲を再生できる）という利点がある。ただし、対比例の構成も本発明の範囲には包含され得る。

[0046] <第2実施形態>

本発明の第2実施形態を説明する。なお、以下に説明する各例示において作用または機能が第1実施形態と同様である要素については、第1実施形態の説明で使用した符号を流用して各々の詳細な説明を適宜に省略する。

[0047] 類似指標 M の算定に利用される編集距離 ε は、歌唱音声が高い（音高情報 Z の総数が多い）ほど大きい数値になるという傾向がある。したがって、歌唱音声の時間長 T が長いにも関わらず編集距離 ε が類似指標 M が小さい数値に維持される候補楽曲は、利用者が歌唱した楽曲に該当する可能性が十分に高いと評価できる。以上の事情を考慮して、第2実施形態の楽曲検索部 62 は、第1処理 SB2において、編集距離 ε を時間長 T で除算した数値を類似指標 M として算定する ($M = \varepsilon / T$)。なお、歌唱音声の時間長 T は、歌唱音声の音符の総数とも換言され得る。N 個の候補楽曲のうち、以上の手順で算定された類似指標 M が、閾値 M_{th} を下回る最小値 M_{min} となる候補楽曲が目標楽曲として確定される点は、第1実施形態と同様である。

[0048] 第2実施形態においても第1実施形態と同様の効果が実現される。また、第2実施形態では、編集距離 ε を歌唱音声の時間長 T で除算した数値が類似指標 M として算定されるから、利用者が歌唱した楽曲である可能性が高い目標楽曲を検索できるという利点がある。

[0049] <変形例>

以上の各形態は多様に変形され得る。具体的な変形の態様を以下に例示する。以下の例示から任意に選択された2以上の態様は適宜に併合され得る。

[0050] (1) 歌唱音声と候補楽曲との間の類似指標 M は、前述の各形態で例示した編集距離 ε に限定されない。例えば、候補楽曲の音高情報 Y の時系列と歌唱音声の音高情報 Z の時系列との間のユークリッド距離または Earth Mover's 距離等の公知の距離尺度を類似指標 M として利用することも可能である。ま

た、歌唱音声と候補楽曲との相関を類似指標Mとして算定することも可能である。歌唱音声と候補楽曲との相関を類似指標Mとして算定する構成では、歌唱音声と候補楽曲とが類似するほど類似指標Mは大きい数値となる。したがって、楽曲検索処理の第2処理SCでは、類似指標Mが最大値Mmaxとなる暫定楽曲が選択され（SC1）、最大値Mmaxが閾値Mthを上回る場合（SC2：YES）に当該暫定楽曲が目標楽曲として確定される。以上の説明から理解される通り、類似指標Mは、歌唱音声と候補楽曲との類似の度合（距離または相関）の指標として包括的に表現され、類似指標Mの算定方法、および、歌唱音声と候補楽曲との類似の度合に対して類似指標Mの数値が増加するか減少するかは不問である。

[0051] （2）前述の各形態では、類似指標Mに応じて選定された1個の目標楽曲を結果通知（S3）として端末装置12の利用者に提示したが、類似指標Mの順番（例えば編集距離 ε 等の距離尺度を類似指標Mとした構成では類似指標Mの昇順）で選択された複数の候補楽曲を結果通知として端末装置12に送信して利用者に提示することも可能である。操作装置28に対する操作で利用者が複数の候補楽曲の何れかを目標楽曲として選択すると、再生制御部64は、当該目標楽曲の楽曲ファイルFを含む再生要求を端末装置12に送信する（S5）。

[0052] 例えば、楽曲検索処理で検索された複数の候補楽曲のリストが端末装置12の表示装置26に表示される。具体的には、複数の候補楽曲の楽曲名を類似指標Mの順番で配列したリストが表示される。類似指標Mに応じて各楽曲の表示態様（例えば表示の色またはサイズ）を相違させることも可能である。利用者は、操作装置28を操作することで、自身が意図した楽曲を目標楽曲としてリストから選択可能である。表示装置26は、利用者が選択した目標楽曲を強調表示する。例えば、利用者が選択した目標楽曲がリストの最上位に移行され、他の楽曲とは異なる表示態様で表示される。以上の例示のように複数の候補楽曲を利用者に提示する構成では、利用者が所望の楽曲を目標楽曲として選択した時点で検索が完了する。すなわち、利用者による目標

楽曲の選択を契機として検索要求の生成および送信が終了し、以降は楽曲検索処理は実行されない。

[0053] (3) 前述の各形態では、利用者が発音した歌唱音声に対応する目標楽曲を検索したが、目標楽曲の検索に利用される音響は歌唱音声に限定されない。例えば、利用者による楽器演奏で発音された演奏音に対応する目標楽曲を楽曲検索部 62 が検索することも可能である。すなわち、收音装置 32 で收音した演奏音を表す音響信号 XA に対して前述の各形態と同様の処理が実行される。以上の説明から理解される通り、利用者が楽曲を発音（歌唱または演奏）した任意の音響（対象音）が目標楽曲の検索に利用される。

[0054] (4) 前述の各形態では、端末装置 12 の音響解析部 42 が音響信号 XA から生成した音高情報 Z の時系列を楽曲処理装置 14 に送信したが、音響解析部 42 を楽曲処理装置 14 に搭載することも可能である。具体的には、收音装置 32 が生成した音響信号 XA が端末装置 12 から通信網 16 を介して楽曲処理装置 14 に送信され、楽曲処理装置 14 の音響解析部 42 が音響信号 XA の解析で音高情報 Z の時系列を生成する。ただし、以上の構成では音響信号 XA を端末装置 12 から楽曲処理装置 14 に送信する必要があるから、通信網 16 の通信負荷を軽減する観点からは、音響解析部 42 を端末装置 12 に搭載して音高情報 Z を楽曲処理装置 14 に送信する構成（すなわち、音響信号 XA を楽曲処理装置 14 に送信する必要がない構成）が有利である。

[0055] (5) 前述の各形態では、候補楽曲と歌唱音声との間で各音符の 1 対 1 の対応を第 1 処理 SB2（音符対応情報 C の生成）にて解析したが、複数の音符の時系列（以下「音符列」という）同士の対応を候補楽曲と歌唱音声との間で解析することも可能である。例えば、時系列に配列した所定個の音符を単位（音符列）として第 1 処理 SB2 を実行することで、候補楽曲と歌唱音声との間における各音符列の対応（音符対応情報 C）が解析される。以上の説明から理解される通り、候補楽曲と歌唱音声（対象音）とについて第 1 処理 SB2 で解析される「各音符の時間的な対応」は、前述の各形態で例示した個々の音符同士の 1 対 1 の対応のほか、複数の音符で構成される音符列同士の対応

も包含する。

[0056] (6) 前述の各形態では、端末装置 12 と通信するサーバ装置を楽曲処理装置 14 として例示したが、楽曲処理装置 14 を単体の装置（例えば携帯電話機およびスマートフォン等の情報端末）で実現することも可能である。例えば、図 10 に例示された楽曲処理装置 14 は、制御装置 50 と記憶装置 52 と收音装置 32 と再生装置 34 とを具備する。記憶装置 52 は、前述の各形態と同様に、楽曲ファイル F と検索情報 V とを N 個の候補楽曲の各々について記憶する。制御装置 50 は、收音装置 32 が生成する音響信号 XA に応じた音高情報 Z を図 2 の音響解析処理で生成する音響解析部 42 と、音高情報 Z を利用した図 4 の楽曲検索処理により目標楽曲を検索する楽曲検索部 62 と、楽曲検索部 62 が検索した目標楽曲を再生装置 34 に再生させる再生制御部 64 とを実現する。具体的には、再生制御部 64 は、目標楽曲の楽曲ファイル F と再生情報 D とを第 1 実施形態と同様に生成し、楽曲ファイル F と再生情報 D とに応じた音響信号 XB を再生装置 34 に供給する。以上の説明から理解される通り、再生制御部 64 は、目標楽曲を再生装置 34 に再生させる要素として包括的に表現される。第 1 実施形態の例示のように再生装置 34 が楽曲処理装置 14 とは別体として設置されるか、図 10 の例示のように再生装置 34 が楽曲処理装置 14 に設置されるかは不問である。また、なお、記憶装置 52（例えばクラウドストレージ）を楽曲処理装置 14 とは別個に設置し、楽曲処理装置 14 が例えば通信網 16 を介して記憶装置 52 の検索情報 V または楽曲ファイル F を参照する構成も採用され得る。

[0057] (7) 前述の各形態で例示した楽曲処理装置 14 は、前述の各形態の例示の通り、制御装置 50 とプログラムとの協働で実現される。本発明の好適な態様に係るプログラムは、歌唱音声および演奏音等の対象音の音符毎の音高に応じた音高情報 Z を利用した楽曲検索処理（図 4）により、対象音に対応する目標楽曲を当該対象音の進行に並行して検索する楽曲検索部 62、および、目標楽曲のうち当該対象音の進行に対応した時点から対象音の進行に並行して目標楽曲を再生装置 34 に再生させる再生制御部 64 としてコンピュー

タを機能させる。以上に例示したプログラムは、コンピュータが読取可能な記録媒体に格納された形態で提供されてコンピュータにインストールされ得る。記録媒体は、例えば非一過性（non-transitory）の記録媒体であり、C D-R O M等の光学式記録媒体（光ディスク）が好例であるが、半導体記録媒体および磁気記録媒体等の公知の任意の形式の記録媒体を包含し得る。なお、「非一過性の記録媒体」とは、一過性の伝搬信号（transitory, propagating signal）を除く全てのコンピュータ読み取り可能な記録媒体を含み、揮発性の記録媒体を除外するものではない。また、通信網を介した配信の形態でプログラムをコンピュータに配信することも可能である。

[0058] （８）本発明は、前述の各形態に係る楽曲処理装置１４の動作方法（楽曲処理方法）としても特定される。例えば、本発明の一態様に係る楽曲処理方法は、歌唱音声および演奏音等の対象音の音符毎の音高に応じた音高情報Ｚを利用した楽曲検索処理（図４）により、対象音に対応する目標楽曲を当該対象音の進行に並行して検索し、目標楽曲のうち当該対象音の進行に対応した時点から対象音の進行に並行して目標楽曲を再生装置３４に再生させる（Ｓ５）。

[0059] （９）以上に例示した具体的な形態から把握される本発明の好適な態様を以下に例示する。

<態様１>

本発明の好適な態様（態様１）に係る楽曲処理方法は、コンピュータシステムが、音高が時間的に変化する対象音の音高に応じた音高情報を利用した楽曲検索処理により、対象音に対応する目標楽曲を当該対象音の進行に並行して検索し、目標楽曲のうち当該対象音の進行に対応した部分から対象音の進行に並行して目標楽曲を再生装置に再生させる。態様１では、対象音に対応する目標楽曲が対象音の進行に並行して検索され、かつ、目標楽曲のうち対象音の進行に対応した部分から当該目標楽曲が当該対象音の進行に並行して再生される。すなわち、楽曲の一連の対象音の進行に並行して目標楽曲の検索と当該目標楽曲の再生とが順次に実行される。したがって、目標楽曲の

検索と当該目標楽曲の再生に並行した対象音の発音（例えば目標楽曲の歌唱または演奏）とを所望する利用者の負担を軽減することが可能である。

[0060] ＜態様２＞

態様１の好適例（態様２）では、目標楽曲の検索において、楽曲検索処理の結果通知から所定時間が経過するまでに利用者から楽曲変更が指示された場合に楽曲検索処理を再実行し、目標楽曲の再生において、楽曲変更が指示されずに所定時間が経過した場合に、直前の楽曲検索処理で検索された目標楽曲を再生装置に再生させる。態様２では、楽曲検索処理の結果通知から所定時間が経過するまでに利用者が楽曲変更を指示した場合に楽曲検索処理が再実行される。したがって、楽曲検索処理による誤検索を修正して適正な目標楽曲を再生することが可能である。

[0061] ＜態様３＞

態様１または態様２の好適例（態様３）において、楽曲検索処理は、複数の候補楽曲の各々について、対象音と当該候補楽曲との間で音高情報の時系列を対比することで、対象音における複数の音符と候補楽曲における複数の音符との時間的な対応と、対象音と当該候補楽曲との類否の指標である類似指標とを生成する第１処理と、各候補楽曲の類似指標に応じて複数の候補楽曲から目標楽曲を選択する第２処理とを含み、目標楽曲の再生においては、対象音と目標楽曲との間で第１処理により解析された音符毎の時間的な対応を参照して、対象音の発音速度に対応する当該目標楽曲の再生速度を設定し、目標楽曲を当該再生速度で再生装置に再生させる。態様３では、対象音の発音速度に対応する再生速度で目標楽曲が再生されるから、楽曲検索処理の実行後に、目標楽曲の再生に並行した対象音の発音を利用者が違和感なく継続できるという利点がある。また、目標楽曲を検索する楽曲検索処理のうち第１処理の結果（すなわち、対象音と目標楽曲との間の各音符の時間的な対応）が、目標楽曲の再生速度の設定に流用される。したがって、目標楽曲の再生速度を楽曲検索処理とは別個の処理で設定する構成と比較して、再生制御部の処理負荷を軽減することが可能である。

[0062] ＜態様４＞

態様３の好適例（態様４）では、目標楽曲の再生において、対象音と目標楽曲との間で第１処理により対応が特定された音符間の音高差に応じて当該目標楽曲の再生調を設定し、目標楽曲を当該再生調で再生装置に再生させる。態様４では、対象音と目標楽曲との間で相互に対応する音符間の音高差に応じた再生調で目標楽曲が再生されるから、楽曲検索処理の実行後に、目標楽曲の再生に並行した対象音の発音を利用者が違和感なく継続できるという利点がある。また、目標楽曲を検索する楽曲検索処理のうち第１処理の結果（すなわち、対象音と目標楽曲との間の各音符の時間的な対応）が、目標楽曲の再生調の設定に流用される。したがって、目標楽曲の再生調を楽曲検索処理とは別個の処理で設定する構成と比較して、再生制御部の処理負荷を軽減することが可能である。

[0063] ＜態様５＞

態様３または態様４の好適例（態様５）では、目標楽曲の再生において、目標楽曲を表す楽曲ファイルと、対象音と目標楽曲との間の時間的な対応を指定する時間対応情報と、目標楽曲の再生速度とを、再生装置を具備する端末装置に通信網を介して送信する。以上の態様では、対象音と目標楽曲との間の時間的な対応を指定する時間対応情報と目標楽曲の再生速度とが目標楽曲の楽曲ファイルとともに端末装置に送信される。したがって、通信網での伝送遅延の大小に関わらず、目標楽曲のうち対象音の進行に応じた部分を特定できる（対象音の進行に同期するように楽曲ファイルの目標楽曲を再生できる）という利点がある。

[0064] ＜態様６＞

本発明の好適な態様（態様６）に係る楽曲処理装置は、音高が時間的に変化する対象音の音高に応じた音高情報を利用した楽曲検索処理により、対象音に対応する目標楽曲を当該対象音の進行に並行して検索する楽曲検索部と、目標楽曲のうち当該対象音の進行に対応した部分から対象音の進行に並行して目標楽曲を再生装置に再生させる再生制御部とを具備する。態様６では

、対象音に対応する目標楽曲が対象音の進行に並行して検索され、かつ、目標楽曲のうち対象音の進行に対応した部分から当該目標楽曲が当該対象音の進行に並行して再生される。すなわち、楽曲の一連の対象音の進行に並行して目標楽曲の検索と当該目標楽曲の再生とが順次に実行される。したがって、目標楽曲の検索と当該目標楽曲の再生に並行した対象音の発音（例えば目標楽曲の歌唱または演奏）とを所望する利用者の負担を軽減することが可能である。

符号の説明

[0065] 1 0 0 ……楽曲処理システム、 1 2 ……端末装置、 1 4 ……楽曲処理装置、 1 6 ……通信網、 2 0, 5 0 ……制御装置、 2 2, 5 2 ……記憶装置、 2 4, 5 4 ……通信装置、 2 6 ……表示装置、 2 8 ……操作装置、 3 2 ……收音装置、 3 4 ……再生装置、 4 2 ……音響解析部、 4 4 ……再生処理部、 5 0 ……制御装置、 5 2 ……記憶装置、 5 4 ……通信装置、 6 2 ……楽曲検索部、 6 4 ……再生制御部。

請求の範囲

[請求項1]

コンピュータシステムが、

音高が時間的に変化する対象音の音高に応じた音高情報を利用した楽曲検索処理により、前記対象音に対応する目標楽曲を当該対象音の進行に並行して検索し、

前記目標楽曲のうち当該対象音の進行に対応した部分から前記対象音の進行に並行して前記目標楽曲を再生装置に再生させる

楽曲処理方法。

[請求項2]

前記目標楽曲の検索においては、前記楽曲検索処理の結果通知から所定時間が経過するまでに利用者から楽曲変更が指示された場合に前記楽曲検索処理を再実行し、

前記目標楽曲の再生においては、前記楽曲変更が指示されずに前記所定時間が経過した場合に、直前の楽曲検索処理で検索された目標楽曲を前記再生装置に再生させる

請求項1の楽曲処理方法。

[請求項3]

前記楽曲検索処理は、

複数の候補楽曲の各々について、前記対象音と当該候補楽曲との間で音高情報の時系列を対比することで、前記対象音における複数の音符と前記候補楽曲における複数の音符との時間的な対応と、前記対象音と当該候補楽曲との類否の指標である類似指標とを生成する第1処理と、

前記各候補楽曲の類似指標に応じて前記複数の候補楽曲から前記目標楽曲を選択する第2処理とを含み、

前記目標楽曲の再生においては、前記対象音と前記目標楽曲との間で前記第1処理により解析された音符毎の時間的な対応を参照して、前記対象音の発音速度に対応する当該目標楽曲の再生速度を設定し、前記目標楽曲を当該再生速度で前記再生装置に再生させる

請求項1または請求項2の楽曲処理方法。

[請求項4] 前記目標楽曲の再生においては、前記対象音と前記目標楽曲との間で前記第1処理により対応が特定された音符間の音高差に応じて当該目標楽曲の再生調を設定し、前記目標楽曲を当該再生調で前記再生装置に再生させる

請求項3の楽曲処理方法。

[請求項5] 前記目標楽曲の再生においては、前記目標楽曲を表す楽曲ファイルと、前記対象音と前記目標楽曲との間の時間的な対応を指定する時間対応情報と、前記目標楽曲の再生速度とを、前記再生装置を具備する端末装置に通信網を介して送信する

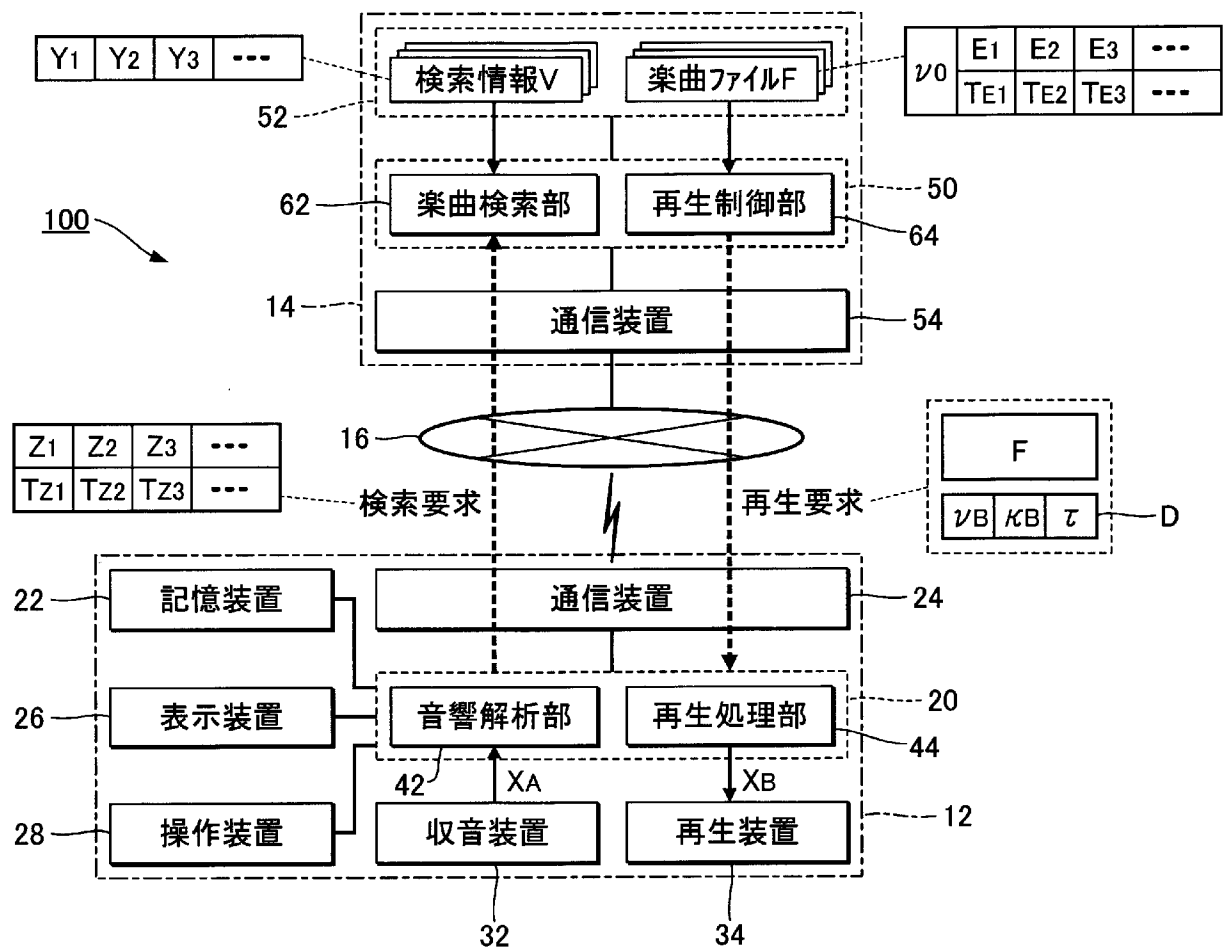
請求項3または請求項4の楽曲処理方法。

[請求項6] 音高が時間的に変化する対象音の音高に応じた音高情報を利用した楽曲検索処理により、前記対象音に対応する目標楽曲を当該対象音の進行に並行して検索する楽曲検索部と、

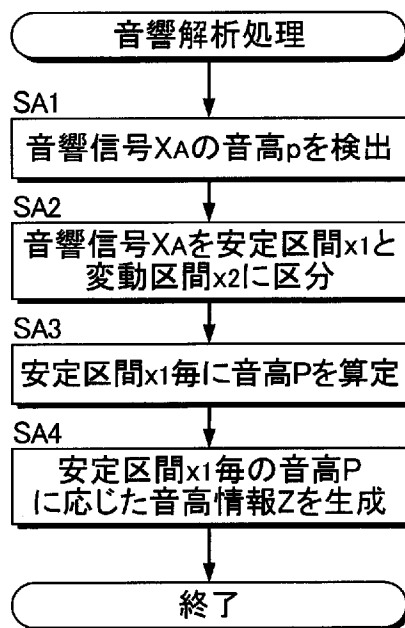
前記目標楽曲のうち当該対象音の進行に対応した部分から前記対象音の進行に並行して前記目標楽曲を再生装置に再生させる再生制御部と

を具備する楽曲処理装置。

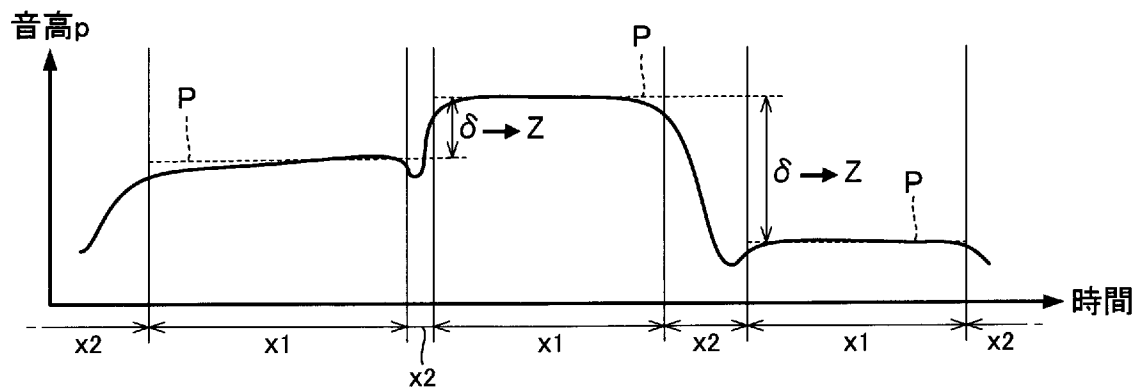
[図1]



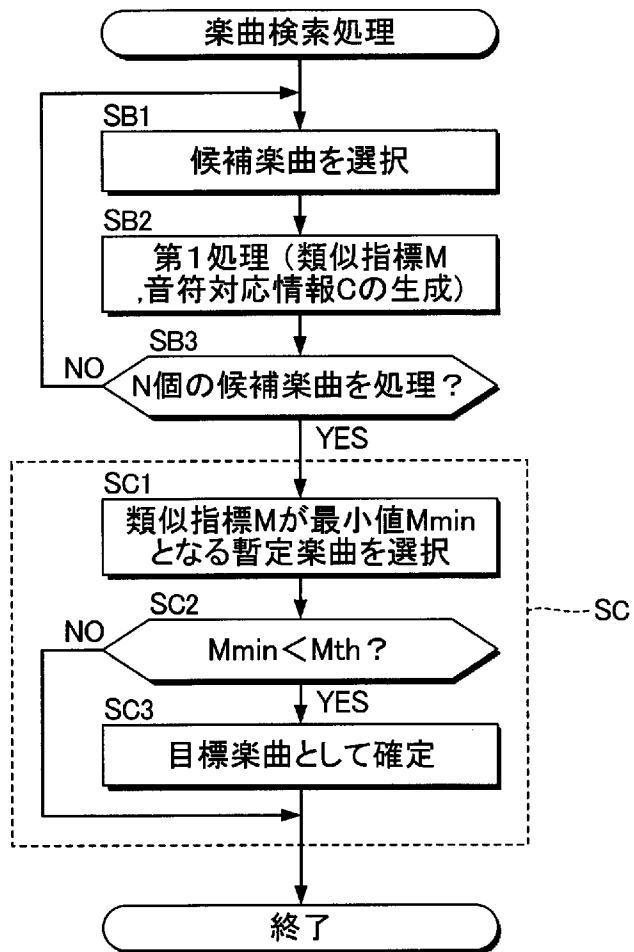
[図2]



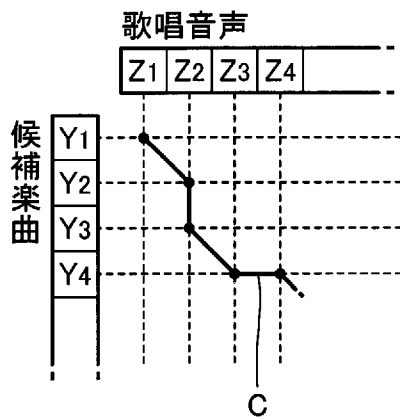
[図3]



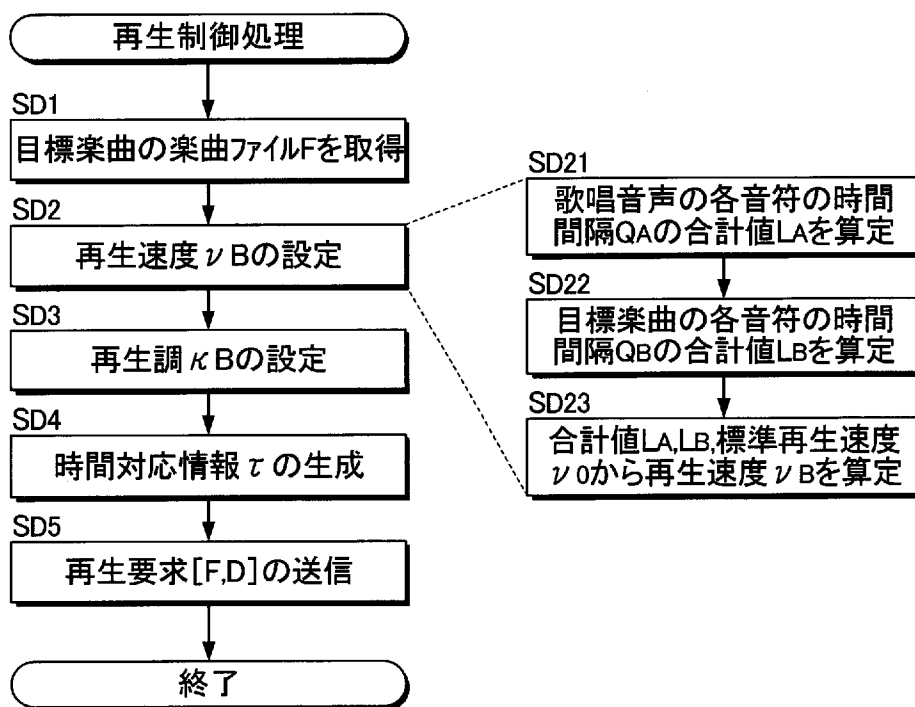
[図4]



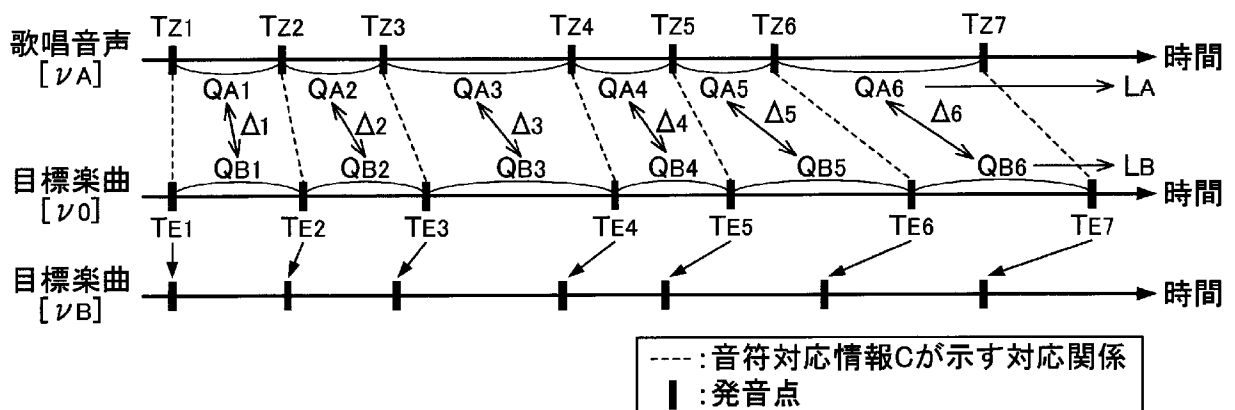
[図5]



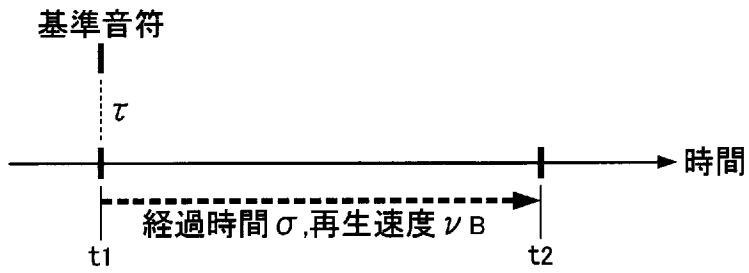
[図6]



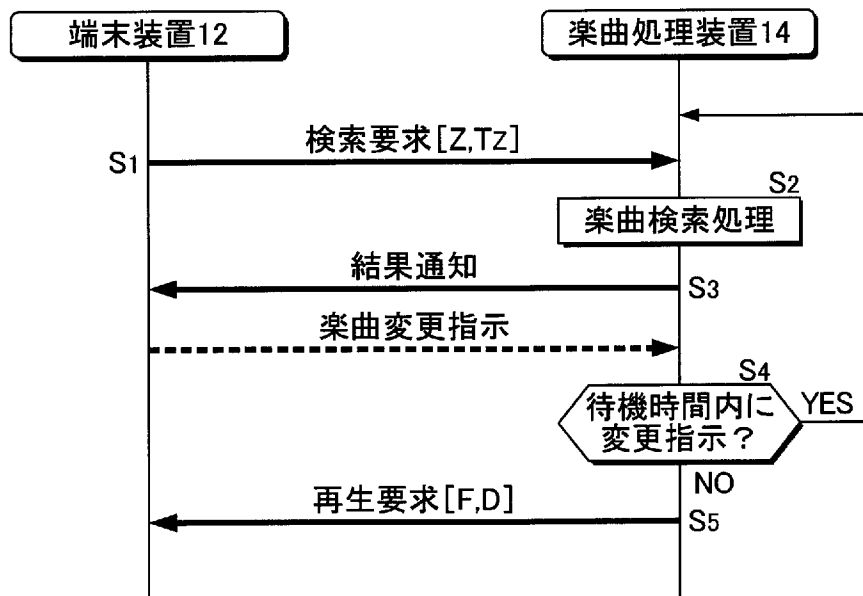
[図7]



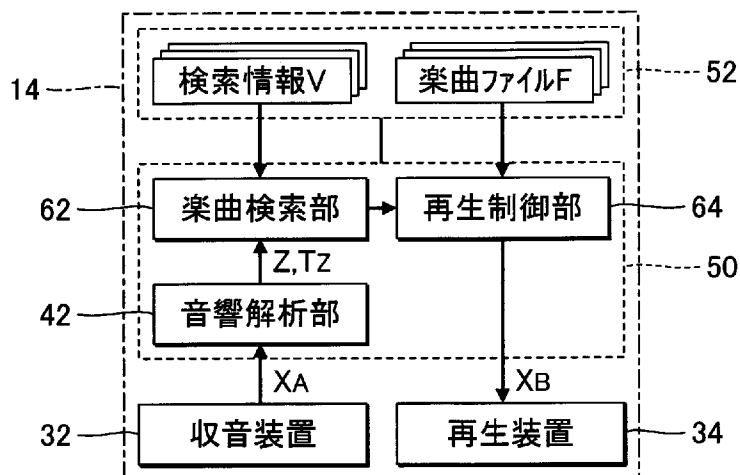
[図8]



[図9]



[図10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/076266

A. CLASSIFICATION OF SUBJECT MATTER

G10L25/54(2013.01)i, G10H1/00(2006.01)i, G10L25/51(2013.01)i, G10L25/90(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L25/54, G10H1/00, G10L25/51, G10L25/90

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	1996-2016
Kokai Jitsuyo Shinan Koho	1971-2016	Toroku Jitsuyo Shinan Koho	1994-2016

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2013-117688 A (Sony Corp.), 13 June 2013 (13.06.2013), paragraphs [0024] to [0065] & US 2014/0318348 A1 & WO 2013/084774 A1 & EP 2790184 A1 paragraphs [0025] to [0066] & CA 2853904 A	1-6
A	JP 7-140977 A (Roland Corp.), 02 June 1995 (02.06.1995), paragraph [0050] (Family: none)	2
A	JP 2002-244672 A (Daiichikoshō Co., Ltd.), 30 August 2002 (30.08.2002), paragraph [0021] (Family: none)	2

☐ Further documents are listed in the continuation of Box C.

☐ 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
14 November 2016 (14.11.16)

Date of mailing of the international search report
22 November 2016 (22.11.16)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G10L25/54(2013.01)i, G10H1/00(2006.01)i, G10L25/51(2013.01)i, G10L25/90(2013.01)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（I P C））

Int.Cl. G10L25/54, G10H1/00, G10L25/51, G10L25/90

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

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 6 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 6 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 6 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

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
X	JP 2013-117688 A（ソニー株式会社）2013.06.13, [0024]-[0065] & US 2014/0318348 A1 & WO 2013/084774 A1 & EP 2790184 A1, [0025]-[0066] & CA 2853904 A	1-6
A	JP 7-140977 A（ローランド株式会社）1995.06.02, [0050]（ファミリーなし）	2
A	JP 2002-244672 A（株式会社第一興商）2002.08.30, [0021]（ファミリーなし）	2

C欄の続きにも文献が列挙されている。

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

1 4 . 1 1 . 2 0 1 6

国際調査報告の発送日

2 2 . 1 1 . 2 0 1 6

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）
郵便番号 1 0 0 - 8 9 1 5
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）

富澤 直樹

5 Z

4 1 8 8

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 1