

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第3879402号
(P3879402)

(45) 発行日 平成19年2月14日(2007.2.14)

(24) 登録日 平成18年11月17日(2006.11.17)

(51) Int.C1.

F 1

G 1 O L 13/02 (2006.01)
G 1 O H 1/00 (2006.01)G 1 O L 13/02 122B
G 1 O H 1/00 1O2Z

請求項の数 5 (全 33 頁)

(21) 出願番号 特願2000-402880 (P2000-402880)
 (22) 出願日 平成12年12月28日 (2000.12.28)
 (65) 公開番号 特開2002-202788 (P2002-202788A)
 (43) 公開日 平成14年7月19日 (2002.7.19)
 審査請求日 平成15年10月23日 (2003.10.23)

前置審査

(73) 特許権者 000004075
 ヤマハ株式会社
 静岡県浜松市中沢町10番1号
 (74) 代理人 100075074
 弁理士 伊沢 敏昭
 (72) 発明者 嘉山 啓
 静岡県浜松市中沢町10番1号ヤマハ株式
 会社内
 (72) 発明者 オスカル セルマン
 スペイン バルセロナ 08002 メル
 セ 12
 (72) 発明者 ジャウメ オルトラ
 スペイン バルセロナ 08002 メル
 セ 12

最終頁に続く

(54) 【発明の名称】歌唱合成方法と装置及び記録媒体

(57) 【特許請求の範囲】

【請求項 1】

子音とこれに続く母音とからなる歌唱音韻について該子音及び母音を表わす音韻情報と該子音及び母音の音高を表わす音高情報と歌唱開始時刻を表わす時刻情報と歌唱長を表わす歌唱長情報とを入力するステップと、

無音又は先行母音から前記子音への音韻遷移時間長を表わし且つ音韻遷移の境界に第1の時間区切り情報を有する第1の時間長情報と前記子音から前記母音への音韻遷移時間長を表わし且つ音韻遷移の境界に第2の時間区切り情報を有する第2の時間長情報とを前記音韻情報及び前記音高情報に基づいて生成するステップと、

前記第1の時間長情報において前記第1の時間区切り情報より後の子音部で表わされる時間長と前記第2の時間長情報において前記第2の時間区切り情報より前の子音部で表わされる時間長とを加算して前記子音の歌唱継続時間を決定すると共に、前記第2の時間長情報において前記第2の時間区切り情報より後の母音部で表わされる時間長に基づき且つ前記歌唱長情報の表わす歌唱長に対応して前記母音の歌唱継続時間を決定する決定ステップであって、前記子音の歌唱開始時刻を前記時刻情報の表わす歌唱開始時刻より前記子音の歌唱継続時間だけ前の時刻に決定し、前記母音の歌唱開始時刻を前記時刻情報の表わす歌唱開始時刻又はその近傍の時刻に決定するものと、

各々前記子音及び母音からなり且つ前記音高情報の表わす音高を有する第1及び第2の歌唱音声を前記決定ステップで決定された前記子音及び母音の歌唱開始時刻にそれぞれ発生開始し、該第1及び第2の歌唱音声を前記決定ステップで決定された前記子音及び母音

の歌唱継続時間中それぞれ発生するステップと
を含む歌唱合成方法。

【請求項 2】

前記決定ステップでは、前記時刻情報の表わす歌唱開始時刻の前後に及ぶ所定の時間範囲内でランダムに新たな歌唱開始時刻を設定し、この設定に係る歌唱開始時刻を前記母音の歌唱開始時刻として決定する請求項 1 記載の歌唱合成方法。

【請求項 3】

子音とこれに続く母音とからなる歌唱音韻について該子音及び母音を表わす音韻情報と該子音及び母音の音高を表わす音高情報と歌唱開始時刻を表わす時刻情報と歌唱長を表わす歌唱長情報とを入力する入力手段と、

無音又は先行母音から前記子音への音韻遷移時間長を表わし且つ音韻遷移の境界に第 1 の時間区切り情報を有する第 1 の時間長情報と前記子音から前記母音への音韻遷移時間長を表わし且つ音韻遷移の境界に第 2 の時間区切り情報を有する第 2 の時間長情報とを記憶する記憶手段と、

前記入力手段で入力された音韻情報及び音高情報に基づいて前記記憶手段から前記第 1 及び第 2 の時間長情報を読み出す読出手段と、

前記読出手段で読み出された第 1 の時間長情報において前記第 1 の時間区切り情報より後の子音部で表わされる時間長と前記読出手段で読み出された第 2 の時間長情報において前記第 2 の時間区切り情報より前の子音部で表わされる時間長とを加算して前記子音の歌唱継続時間を算出すると共に、前記読出手段で読み出された第 2 の時間長情報において前記第 2 の時間区切り情報より後の母音部で表わされる時間長に基づき且つ前記歌唱長情報の表わす歌唱長に対応して前記母音の歌唱継続時間を算出する計算手段であって、前記子音の歌唱開始時刻としては前記時刻情報の表わす歌唱開始時刻より前記子音の歌唱継続時間だけ前の時刻を算出し、前記母音の歌唱開始時刻としては前記時刻情報の表わす歌唱開始時刻又はその近傍の時刻を算出するものと、

各々前記子音及び母音からなり且つ前記音高情報を有する第 1 及び第 2 の歌唱音声を前記計算手段で算出された前記子音及び母音の歌唱開始時刻にそれぞれ発生開始し、該第 1 及び第 2 の歌唱音声を前記計算手段で算出された前記子音及び母音の歌唱継続時間中それぞれ発生する歌唱音声合成手段と

を備えた歌唱合成装置。

【請求項 4】

前記入力手段では、前記第 1 の時間長情報において前記第 1 の時間区切り情報より後の子音部で表わされる時間長と前記第 2 の時間長情報において前記第 2 の時間区切り情報より前の子音部で表わされる時間長とを修正するための修正情報を入力し、前記計算手段では、前記読出手段で読み出された第 1 の時間長情報の前記子音部で表わされる時間長と前記読出手段で読み出された第 2 の時間長情報の前記子音部で表わされる時間長とを前記入力手段で入力された修正情報に応じて修正した後、その修正に係る時間長に基づいて前記子音の歌唱継続時間及び歌唱開始時刻を算出する請求項 3 記載の歌唱合成装置。

【請求項 5】

子音とこれに続く母音とからなる歌唱音韻について該子音及び母音を表わす音韻情報と該子音及び母音の音高を表わす音高情報と歌唱開始時刻を表わす時刻情報と歌唱長を表わす歌唱長情報とを入力するステップと、

無音又は先行母音から前記子音への音韻遷移時間長を表わし且つ音韻遷移の境界に第 1 の時間区切り情報を有する第 1 の時間長情報と前記子音から前記母音への音韻遷移時間長を表わし且つ音韻遷移の境界に第 2 の時間区切り情報を有する第 2 の時間長情報とを前記音韻情報及び前記音高情報に基づいて生成するステップと、

前記第 1 の時間長情報において前記第 1 の時間区切り情報より後の子音部で表わされる時間長と前記第 2 の時間長情報において前記第 2 の時間区切り情報より前の子音部で表わされる時間長とを加算して前記子音の歌唱継続時間を決定すると共に、前記第 2 の時間長情報において前記第 2 の時間区切り情報より後の母音部で表わされる時間長に基づき且つ

10

20

30

40

50

前記歌唱長情報の表わす歌唱長に対応して前記母音の歌唱継続時間を決定する決定ステップであって、前記子音の歌唱開始時刻を前記時刻情報の表わす歌唱開示時刻より前記子音の歌唱継続時間だけ前の時刻に決定し、前記母音の歌唱開始時刻を前記時刻情報の表わす歌唱開始時刻又はその近傍の時刻に決定するものと、

各々前記子音及び母音からなり且つ前記音高情報の表わす音高を有する第1及び第2の歌唱音声を前記決定ステップで決定された前記子音及び母音の歌唱開始時刻にそれぞれ発生開始し、該第1及び第2の歌唱音声を前記決定ステップで決定された前記子音及び母音の歌唱継続時間中それぞれ発生するステップと

をコンピュータに実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

10

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、リアルタイムに入力される演奏データに基づいて歌唱音声を合成する歌唱合成方法と装置及び記録媒体に関するものである。

【0002】

【従来の技術】

従来、この種の歌唱合成装置としては、演奏データのノートオン信号に伴って発生される先頭音素の立上り時間を、この先頭音素がノートオン期間中に他の音素に引き続いて発音される場合の立上り時間より短くする（速く立上らせる）ようにしたもののが知られている（例えば、特開平10-49169号公報参照）。

20

【0003】

図40(A)は、人歌唱における子音歌唱開始時刻及び母音歌唱開始時刻を示すもので、この例では、音高「C₃（ド）」、「D₃（レ）」、「E₃（ミ）」にそれぞれ対応して歌詞「さ（sa）」、「い（i）」、「た（ta）」と歌った場合を示している。図40(A)によれば、「sa」、「ta」のように子音と母音との組合せからなる音韻は、子音が母音に先行して発音開始されているのがわかる。

【0004】

一方、図40(B)は、上記した従来の歌唱合成における歌唱開始時刻を示すもので、歌唱内容は、図40(A)の場合と同様である。実歌唱開始時刻T1～T3は、いずれもノートオン信号による歌唱音声の発生開始時刻に相当する。上記した従来の歌唱合成では、例えば「sa」の歌唱音声を発生する場合、子音「s」の歌唱開始時刻を時刻T1に合わせ、時刻T1から子音「s」の振幅レベルを急速に上昇させることにより楽器音（伴奏音）に対して歌唱音声が遅れた感じとなる事態を回避するようにしている。

30

【0005】

【発明が解決しようとする課題】

上記した従来の歌唱合成技術によると、次のような問題点がある。

【0006】

(1) 図40(A)の人歌唱において母音歌唱開始時刻は、図40(B)の歌唱合成における実歌唱開始時刻（ノートオン時刻）にほぼ対応する。図40(B)の場合は、子音歌唱開始時刻をノートオン時刻に一致させると共に子音（先頭音素）の立上り時間を短縮しているので、図40(A)の場合に比べて歌唱開始時刻や歌唱継続時間が不自然となる。

40

【0007】

(2) ノートオン時刻の直前に音韻情報を送信し、ノートオン時刻に音韻情報対応の歌唱音声の発生を開始するようとしているので、ノートオン時刻より早く歌唱音声を発生開始することができない。

【0009】

この発明の目的は、上記のような問題点を解決した新規な歌唱合成方法と装置を提供することにある。

【0010】

50

【課題を解決するための手段】

この発明に係る歌唱合成方法は、

子音とこれに続く母音とからなる歌唱音韻について該子音及び母音を表わす音韻情報と該子音及び母音の音高を表わす音高情報と歌唱開始時刻を表わす時刻情報と歌唱長を表わす歌唱長情報とを入力するステップ(図4)と、

無音又は先行母音から前記子音への音韻遷移時間長を表わし且つ音韻遷移の境界に第1の時間区切り情報を有する第1の時間長情報と前記子音から前記母音への音韻遷移時間長を表わし且つ音韻遷移の境界に第2の時間区切り情報を有する第2の時間長情報とを前記音韻情報及び前記音高情報を基づいて生成するステップ(0066, 図23)と、

前記第1の時間長情報において前記第1の時間区切り情報を基づいて前記子音部で表わされる時間長と前記第2の時間長情報において前記第2の時間区切り情報を基づいて前記母音部で表わされる時間長とを加算して前記子音の歌唱継続時間を決定すると共に、前記第2の時間長情報において前記第2の時間区切り情報を基づいて前記母音部で表わされる時間長に基づき且つ前記歌唱長情報を基づいて前記母音の歌唱継続時間を決定する決定ステップであって、前記子音の歌唱開始時刻を前記時刻情報の表わす歌唱開始時刻より前記子音の歌唱継続時間だけ前の時刻に決定し、前記母音の歌唱開始時刻を前記時刻情報の表わす歌唱開始時刻又はその近傍の時刻に決定するもの(0093, 図19~33)と、

各々前記子音及び母音からなり且つ前記音高情報を基づいて第1及び第2の歌唱音声を前記決定ステップで決定された前記子音及び母音の歌唱開始時刻にそれぞれ発生開始し、該第1及び第2の歌唱音声を前記決定ステップで決定された前記子音及び母音の歌唱継続時間中それぞれ発生するステップ(図13)とを含むものである。

【0011】

この発明の歌唱合成方法によれば、音韻情報及び音高情報を基づいて第1及び第2の時間長情報を生成した上で第1の時間長情報の子音部の時間長と第2の時間長情報の母音部の時間長とを加算して子音の歌唱継続時間を決定すると共に第2の時間長情報の母音部の時間長に基づき且つ歌唱長情報を基づいて母音の歌唱継続時間を決定し、子音の歌唱開始時刻を時刻情報の表わす歌唱開始時刻より子音の歌唱継続時間だけ前の時刻に決定し、母音の歌唱開始時刻を時刻情報の表わす歌唱開始時刻又はその近傍の時刻に決定するようにしたので、各々子音及び母音からなる第1及び第2の歌唱音声として自然な歌唱音声を発生することができ、しかも子音の立上りを母音の立上りより十分に早くして人歌唱に近似した自然な歌唱合成を行なうことができる。この発明の歌唱合成方法において、前記決定ステップでは、前記時刻情報の表わす歌唱開始時刻の前後に及ぶ所定の時間範囲内でランダムに新たな歌唱開始時刻を設定し、この設定に係る歌唱開始時刻を前記母音の歌唱開始時刻として決定するようにしてもよい(0093)。このようにすると、母音の歌唱開始タイミングに変化がある自然な歌唱合成が可能となる。

【0014】

この発明に係る歌唱合成装置は、

子音とこれに続く母音とからなる歌唱音韻について該子音及び母音を表わす音韻情報と該子音及び母音の音高を表わす音高情報と歌唱開始時刻を表わす時刻情報と歌唱長を表わす歌唱長情報とを入力する入力手段(図2の30, 図4)と、

無音又は先行母音から前記子音への音韻遷移時間長を表わし且つ音韻遷移の境界に第1の時間区切り情報を有する第1の時間長情報と前記子音から前記母音への音韻遷移時間長を表わし且つ音韻遷移の境界に第2の時間区切り情報を有する第2の時間長情報とを記憶する記憶手段(0066, 図2の14)と、

前記入力手段で入力された音韻情報及び音高情報を基づいて前記記憶手段から前記第1及び第2の時間長情報を読み出す読出手段(図3の42)と、

前記読出手段で読み出された第1の時間長情報を基づいて前記第1の時間区切り情報を基づいて前記子音部で表わされる時間長と前記読出手段で読み出された第2の時間長情報を基づいて前記第2の時間区切り情報を基づいて前記母音部で表わされる時間長とを加算して前記子音の歌

10

20

30

40

50

唱継続時間を算出すると共に、前記読み出手段で読み出された第2の時間長情報において前記第2の時間区切り情報より後の母音部で表わされる時間長に基づき且つ前記歌唱長情報の表わす歌唱長に対応して前記母音の歌唱継続時間を算出する計算手段であって、前記子音の歌唱開始時刻としては前記時刻情報の表わす歌唱開始時刻より前記子音の歌唱継続時間だけ前の時刻を算出し、前記母音の歌唱開始時刻としては前記時刻情報の表わす歌唱開始時刻又はその近傍の時刻を算出するもの(0093, 図19~33)と、

各々前記子音及び母音からなり且つ前記音高情報の表わす音高を有する第1及び第2の歌唱音声を前記計算手段で算出された前記子音及び母音の歌唱開始時刻にそれぞれ発生開始し、該第1及び第2の歌唱音声を前記計算手段で算出された前記子音及び母音の歌唱継続時間中それぞれ発生する歌唱音声合成手段(図2の28, 図13)と
10を備えたものである。

【0015】

この発明の歌唱合成装置は、この発明の歌唱合成方法を実施するものであり、この発明の歌唱合成方法に関して前述したと同様の作用効果が得られる。また、音韻遷移時間長を記憶手段から読み出して利用するようにしたので、歌唱音韻数が増大しても、構成乃至処理が簡単である。

【0016】

この発明の歌唱合成装置において、前記入力手段では、前記第1の時間長情報において前記第1の時間区切り情報より後の子音部で表わされる時間長と前記第2の時間長情報において前記第2の時間区切り情報より前の子音部で表わされる時間長とを修正するための修正情報を入力し、前記計算手段では、前記読み出手段で読み出された第1の時間長情報の前記子音部で表わされる時間長と前記読み出手段で読み出された第2の時間長情報の前記子音部で表わされる時間長とを前記入力手段で入力された修正情報に応じて修正した後、その修正に係る時間長に基づいて前記子音の歌唱継続時間及び歌唱開始時刻を算出するようにもよい(図25, 26, 29, 30)。このようにすると、子音の歌唱継続時間及び歌唱開始時刻に入力側の意志を反映させることができ、一層自然な歌唱合成が可能となる。
20

【0025】

【発明の実施の形態】

図1を参照してこの発明の歌唱合成の概要を説明する。図1(A)は、図40(A)と同様に人歌唱における子音歌唱開始時刻及び母音歌唱開始時刻を示すものであり、図1(B)は、この発明の歌唱合成における歌唱開始時刻を示すものである。
30

【0026】

この発明では、演奏データは、歌詞を構成する「s a」等の音韻毎に音韻情報、歌唱開始時刻情報、歌唱長情報等を入力する。歌唱開始時刻情報は、T1等の実歌唱開始時刻(例えば拍頭タイミング)を表わす。各演奏データは、実歌唱開始時刻より早く入力し、各演奏データ毎に音韻情報を音韻遷移時間長に変換する。音韻遷移時間長は、例えば第1の音素「s」及び第2の音素「a」からなる音韻「s a」については第1の音素の発生時間長と第2の音素の発生時間長とからなるもので、この音韻遷移時間長と歌唱開始時刻情報と歌唱長情報とを用いて第1及び第2の音素の歌唱開始時刻と歌唱継続時間とを決定する。
40 このとき、子音「s」の歌唱開始時刻は、実歌唱開始時刻T1より前とし、母音「a」の歌唱開始時刻は、実歌唱開始時刻T1と同一あるいはその前又は後とすることができます。このことは、「i」と「t a」についても同様である。図1(B)の例では、「s a」について「s」の歌唱開始時刻を(A)の人歌唱に合わせてT1より前とし、「a」の歌唱開始時刻をT1とし、「i」の歌唱開始時刻をT2とし、「t a」について「t」の歌唱開始時刻を(A)の人歌唱に合わせてT3より前とし、「a」の歌唱開始時刻をT3としている。

【0027】

歌唱合成では、例えば「s」を決定に係る歌唱開始時刻で発生開始し、決定に係る歌唱継続期間中発生する。また、「a」を決定に係る歌唱開始時刻で発生開始し、決定に係る歌
50

唱継続期間中発生する。このことは、「i」及び「t a」についても同様である。この結果、図1(A)の人歌唱の場合と歌唱開始時刻や歌唱継続時間が近似した自然性豊かな歌唱合成が可能となる。

【0028】

図2は、この発明の一実施形態に係る歌唱合成装置の回路構成を示すものである。この歌唱合成装置は、小型コンピュータによって動作が制御されるものである。

【0029】

バス10には、CPU(中央処理装置)12、ROM(リード・オンリィ・メモリ)14、RAM(ランダム・アクセス・メモリ)16、検出回路20、表示回路22、外部記憶装置24、タイマ26、音源回路28、MIDI(Musical Instrument Digital Interface)インターフェース30等が接続されている。10

【0030】

CPU12は、ROM14にストアされたプログラムに従って楽音発生、歌唱合成等に関する各種処理を実行するもので、歌唱合成に関する処理については図17等を参照して後述する。

【0031】

RAM16は、CPU12の各種処理に際してワーキングエリアとして使用される種々の記憶部を含むもので、この発明の実施に關係する記憶部としては、受信した演奏データを書込むための受信バッファ等が設けられている。

【0032】

検出回路20は、パネルの操作子群34から各種操作子の操作情報を検出するものである。20

【0033】

表示回路22は、表示器36の表示動作を制御することにより各種の表示を可能にするものである。

【0034】

外部記憶装置24は、HD(ハードディスク)、FD(フロッピーディスク)、CD(コンパクトディスク)、DVD(デジタル多目的ディスク)、MO(光磁気ディスク)等のうち1又は複数種類の記録媒体を着脱可能なものである。外部記憶装置24に所望の記録媒体を装着した状態では、記録媒体からRAM16へデータを転送可能である。また、装着した記録媒体がHDやFDのように書き込み可能なものであれば、RAM16のデータを記録媒体に転送可能である。30

【0035】

プログラム記録手段としては、ROM14の代わりに外部記憶装置24の記録媒体を用いることができる。この場合、記録媒体に記録したプログラムは、外部記憶装置24からRAM16へ転送する。そして、RAM16に記憶したプログラムに従ってCPU12を動作させる。このようにすると、プログラムの追加やバージョンアップ等を容易に行なうことができる。

【0036】

タイマ26は、テンポデータTMの指示するテンポに対応した周期でテンポクロック信号TCLを発生するもので、テンポクロック信号TCLは、CPU12に割り込み命令信号として供給される。CPU12は、テンポクロック信号TCLに基づく割り込み処理により歌唱合成を行なう。テンポデータTMの指示するテンポは、操作子群34中のテンポ設定操作子の操作に応じて可変設定することができる。テンポクロック信号TCLの発生周期は、一例として5[m s]とすることができる。40

【0037】

音源回路28は、多数の楽音発生チャンネル及び多数の歌唱合成チャンネルを含むもので、歌唱合成チャンネルは、フォルマント合成方式により歌唱音声を合成するようになっている。後述する歌唱合成処理では、歌唱合成チャンネルから歌唱音声信号が発生される。発生に係る楽音信号及び/または歌唱音声信号は、サウンドシステム38により音響に変

換される。

【0038】

MIDIインターフェース30は、この歌唱合成装置とは別体のMIDI機器39との間でMIDI通信を行なうために設けられたもので、この発明では、MIDI機器39から歌唱合成用の演奏データを受信するために用いられる。MIDI機器39からは、歌唱合成用の演奏データと共に歌唱伴奏用の演奏データを受信し、歌唱伴奏用の演奏データに基づいて音源回路28で伴奏用の楽音信号を発生し、サウンドシステム38から伴奏音を発生させるようにしてもよい。

【0039】

次に、図3を参照して実施形態に係る歌唱合成処理を概略的に説明する。ステップ40では、演奏データを入力する。すなわち、MIDI機器39からMIDIインターフェース30を介して演奏データを受信する。演奏データの詳細については、図4を参照して後述する。

【0040】

ステップ42では、受信した演奏データ毎に該演奏データに基づいて歌唱合成DB(データベース)14A内の音韻遷移DB14b及び状態遷移DB14cからそれぞれ音韻遷移時間長及び状態遷移時間長を獲得し、これらの音韻遷移時間長及び状態遷移時間長と演奏データとに基づいて歌唱合成スコアを形成する。歌唱合成スコアは、音韻トラックと、遷移トラックと、ビブラートトラックとの3トラックを有するもので、音韻トラックは、歌唱開始時刻及び歌唱継続時間等の情報を含み、遷移トラックは、アタック等の遷移状態の開始時刻及び継続時間等の情報を含み、ビブラートトラックは、ビブラート付加状態の開始時刻及び継続時間等の情報を含む。

【0041】

ステップ44では、歌唱合成エンジンによる歌唱合成を行なう。すなわち、ステップ40で入力した演奏データとステップ42で形成した歌唱合成スコアと歌唱合成DB14A内の音韻DB14a、音韻遷移DB14b、状態遷移DB14c及びビブラートDB14dから獲得した音源制御情報とに基づいて歌唱合成を行ない、歌唱順に歌唱音声信号を発生する。歌唱合成処理では、音韻トラックの指示する单一音韻(例えば「a」)又は遷移する音韻(例えば「s」から「a」に遷移する「sa」)からなり且つ演奏データの指示する音高を有する歌唱音声を音韻トラックの指示する歌唱開始時刻に発生開始し、音韻トラックの指示する歌唱継続時間中発生する。

【0042】

このようにして発生される歌唱音声には、遷移トラックの指示するアタック等の遷移状態の開始時刻から音高、振幅等の微小変化を付加し、この付加状態を遷移トラックの指示するアタック等の遷移状態の継続時間中継続させることができる。また、歌唱音声には、ビブラートトラックの指示する開始時刻からビブラートを付加し、この付加状態をビブラートトラックの指示する継続時間中継続させることができる。

【0043】

ステップ46, 48は、音源回路28内の処理を示したもので、ステップ46では、歌唱音声信号をD/A(デジタル/アナログ)変換し、ステップ48では、D/A変換された歌唱音声信号をサウンドシステム38に出力し、歌唱音として発音させる。

【0044】

図4は、演奏データ内の情報を示すものである。演奏データは、1音節を歌唱するために必要な演奏情報を含んでおり、演奏情報は、ノート情報と、音韻トラック情報と、遷移トラック情報と、ビブラートトラック情報とを含んでいる。

【0045】

ノート情報には、実歌唱開始時刻を表わすノートオン(Note On)情報、実歌唱長を表わすデュレーション(Duration)情報、歌唱音高を表わすピッチ(Pitch)情報等が含まれる。音韻トラック情報には、歌唱音韻を表わす音韻(PhUと表記する)情報、歌唱子音伸縮率を表わす子音修正(Consonant Modification)情報等が含まれる。この実施形態で

10

20

30

40

50

は、日本語の歌唱合成を行なうことを前提としているので、出現する音素は、子音と母音であり、音韻状態(PhU State)としては、子音と母音との組合せ、母音のみ又は有声子音(鼻音、半母音)のみのいずれかとなる。ここで、有声子音のみの場合は、有声子音の歌唱開始タイミングが母音のみの場合と類似しているので、母音のみの場合として扱うこととする。

【0046】

遷移トラック情報には、歌唱アタックタイプを表わすアタックタイプ(Attack Type)情報、歌唱アタック伸縮率を表わすアタックレート(Attack Rate)情報、歌唱リリースタイプを表わすリリースタイプ(Release Type)情報、歌唱リリース伸縮率を表わすリリースレート(Release Rate)情報、歌唱ノート遷移タイプを表わすノート遷移タイプ(Note Transition Type)情報、歌唱ノート遷移伸縮率を表わすノート遷移レート(Note Transition Rate)情報等が含まれる。アタックタイプ情報により指定されるアタックタイプとしては、「ノーマル(Normal)」、「セクシー(Sexy)」、「シャープ(Sharp)」、「ソフト(Soft)」等がある。リリースタイプ情報及びノート遷移タイプ情報も、アタックタイプと同様の複数のタイプのうちのいずれかを指定可能である。ノート遷移(Note Transition)は、本演奏データ(演奏イベント)から次の演奏データ(演奏イベント)への遷移を意味する。歌唱アタック伸縮率、歌唱リリース伸縮率及び歌唱ノート遷移伸縮率は、状態遷移時間長を伸長したい場合は1より大きな値とし、状態遷移時間長を短縮したい場合は1より小さな値とする。歌唱アタック伸縮率、歌唱リリース伸縮率及び歌唱ノート遷移伸縮率は、0とすることもでき、この場合はそれぞれアタック、リリース及びノート遷移に伴う音高、振幅等の微小変化の付加は行なわれない。10

【0047】

ビブラートトラック情報には、本演奏データ中のビブラートイベント数を表わすビブラートナンバ(Vibrato Number)情報、1番目のビブラートの遅延時間を表わすビブラート遅延1(Vibrato Delay 1)の情報、1番目のビブラートの継続時間を表わすビブラートデュレーション1(Vibrato Duration 1)の情報、1番目のビブラートのタイプを表わすビブラートタイプ1(Vibrato Type 1)の情報…K(2以上の整数)番目のビブラートの遅延時間を表わすビブラート遅延K(Vibrato Delay K)の情報、K番目のビブラートの継続時間を表わすビブラートデュレーションK(Vibrato Duration K)の情報、K番目のビブラートのタイプを表わすビブラートタイプK(Vibrato Type K)の情報等が含まれる。ビブラートイベント数が0の場合は、ビブラート遅延1の情報以下の情報は含まれない。ビブラートタイプ1～Kの情報によって指定されるビブラートタイプとしては、「ノーマル(Normal)」、「セクシー(Sexy)」、「エンカ(Enka)」等がある。20

【0048】

図3に示した歌唱合成DB14Aは、ROM14内に含まれるものであるが、外部記憶装置24に設けたものをRAM16に転送して利用するようにしてもよい。歌唱合成DB14A内には、音韻DB14a、音韻遷移DB14b、状態遷移DB14c、ビブラートDB14d…その他のDB14nが設けられている。

【0049】

次に、図5～8を参照して音韻DB14a、音韻遷移DB14b、状態遷移DB14c及びビブラートDB14d内の記憶情報について説明する。音韻DB14a及びビブラートDB14dには、それぞれ図5及び図8に示すように音源制御情報が記憶される。音韻遷移DB14bには、図6(B)に示すように音韻遷移時間長及び音源制御情報が記憶されると共に、状態遷移DB14cには、図7に示したように状態遷移時間長及び音源制御情報が記憶される。これらの記憶情報を作成する際には、歌唱者による歌唱音声を分析して音源制御情報、音韻遷移時間長及び状態遷移時間長を決定した。また、「(Normal)」、「(Sexy)」、「Soft」、「Enka」等のタイプについては、歌唱者に同意の教示(例えば「セクシーなアタックを付けて歌って下さい」、「演歌調のビブラートを付けて歌って下さい」等)を与えて歌唱してもらい、その歌唱音声を分析して音源制御情報、音韻遷移時4050

間長及び状態遷移時間長を決定した。音源制御情報は、音源回路28において所望の歌唱音声を合成するために必要なフォルマント周波数及びフォルマントレベルの制御パラメータからなるものである。

【0050】

図5に示す音韻DB14aでは、「a」、「i」、「M」、「Sil」等の音韻毎に且つ1音韻内で「P1」、「P2」…等の音高毎に音源制御情報が記憶される。図5～8及び以下の説明において、「M」は、音韻「u」を表わし、「Sil」は、無音(Silence)を表わす。歌唱合成の際には、合成すべき歌唱音声の音韻と音高とに適合した音源制御情報が音韻DB14aから選択される。

【0051】

図6(A)は、音韻遷移DB14bに記憶される音韻遷移時間長(a)～(f)を示すものである。図6(A)及び以下の説明において、「V_Sil」等の記号の意味は次の通りである。

【0052】

(a)「V_Sil」は、母音(Vowel)から無音(Silence)への音韻遷移を表わし、図6(B)では、例えば先行音韻「a」と後続音韻「Sil」との組合せに対応する。

【0053】

(b)「Sil_C」は、無音(Silence)から子音(Consonant)への音韻遷移を表わし、図6(B)では、例えば先行音韻「Sil」と後続音韻「s」(図示せず)との組合せに対応する。

【0054】

(c)「C_V」は、子音(Consonant)から母音(Vowel)への音韻遷移を表わし、図6(B)では、例えば先行音韻「s」(図示せず)と後続音韻「a」(図示せず)との組合せに対応する。

【0055】

(d)「Sil_V」は、無音(Silence)から母音(Vowel)への音韻遷移を表わし、図6(B)では、例えば先行音韻「Sil」と後続音韻「a」との組合せに対応する。

【0056】

(e)「p V_C」は、先行母音(Vowel)から子音(Consonant)への音韻遷移を表わし、図6(B)では、例えば先行音韻「a」と後続音韻「s」(図示せず)との組合せに対応する。

【0057】

(f)「p V_V」は、先行母音(Vowel)から母音(Vowel)への音韻遷移を表わし、図6(B)では、例えば先行音韻「a」と後続音韻「i」との組合せに対応する。

【0058】

図6(B)に示す音韻遷移DB14bでは、「a」～「i」等の音韻組合せ(遷移する音韻)毎に且つ1つの音韻組合せ内で「P1」、「P2」…等の音高毎に音韻遷移時間長及び音源制御情報が記憶される。図6(B)において、「Aspiration」は、呼気音を示す。音韻遷移時間長は、先行音韻の時間長と後続音韻の時間長との組合せからなるものであり、両時間長の境界を時間区切り情報として保持している。歌唱合成スコアを形成する際には、音韻トラックを形成すべき音韻組合せと音高とに適合した音韻遷移時間長が音韻遷移DB14bから選択される。また、歌唱合成の際には、合成すべき歌唱音声の音韻組合せと音高とに適合した音源制御情報が音韻遷移DB14bから選択される。

【0059】

図7に示す状態遷移DB14cでは、アタック(Attack)、ノート遷移(Note Transition、「NtN」と表記する)及びリリース(Release)の各遷移状態毎に且つ1つの遷移状態内で「Normal」、「Sexy」、「Sharp」及び「Soft」の各状態タイプ毎に状態遷移時間長及び音源制御情報が記憶され、しかも1つの状態タイプ内で「a」、「i」等の音韻毎に且つ1音韻内で「P1」、「P2」等の音高毎に状態遷移時間長及び音源制御情報が記憶される。状態遷移時間長は、アタック、ノート遷移又はリリースの遷移状態の継続時間に

10

20

40

50

相当するものである。歌唱合成スコアを形成する際には、遷移トラックを形成すべき遷移状態と状態タイプと音韻と音高とに適合した状態遷移時間長が状態遷移 D B 1 4 c から選択される。また、歌唱合成の際には、合成すべき歌唱音声の遷移状態と状態タイプと音韻と音高とに適合した音源制御情報が状態遷移 D B 1 4 c から選択される。

【0060】

図8に示すビブラートD B 1 4 dでは、「Normal」、「Sexy」…「Enka」等の各ビブラートタイプ毎に音源制御情報が記憶されると共に、1つのビブラートタイプ内で「a」、「i」等の音韻毎に且つ1音韻内で「P 1」、「P 2」等の音高毎に音源制御情報が記憶される。歌唱合成の際には、合成すべき歌唱音声のビブラートタイプと音韻と音高とに適合した音源制御情報がビブラートD B 1 4 dから選択される。

10

【0061】

図9は、演奏データに基づく歌唱合成を示すものである。演奏データS₁、S₂、S₃が図1(B)で示したと同様に「さ(s a) : C₃ : T1 ...」、「い(i) : D₃ : T2 ...」、「た(t a) : E₃ : T3 ...」をそれぞれ指示するものとすると、演奏データS₁、S₂、S₃は、実歌唱開始時刻T1、T2、T3よりそれぞれ早い時刻t₁、t₂、t₃に送信され、MIDIインターフェース30を介して受信される。演奏データの送受信処理は、ステップ40の演奏データ入力処理に相当する。演奏データが受信されると、ステップ42で演奏データ毎に歌唱合成スコアが形成される。そして、ステップ44では、形成された歌唱合成スコアに従って歌声音声SS₁、SS₂、SS₃が合成される。歌唱合成では、歌唱音声SS₁の子音「s」をT1より早い時刻T₁₁に発生開始させると共に、SS₁の母音「a」をT1に発生開始させることができる。また、歌唱音声SS₂の母音「i」をT2に発生開始させることができる。さらに、歌唱音声SS₃の子音「t」をT3より早く時刻T₃₁に発生開始させると共にSS₃の母音「a」をT3に発生開始させることができる。所望により、「さ」の母音「a」又は「い」の母音「i」をT1又はT2よりそれぞれ早く発生開始させることも可能である。

20

【0062】

図10は、ステップ42における参照スコア及び歌唱合成スコアの形成状況を示すものである。この実施形態では、歌唱合成スコア形成処理の前処理として、参照スコア形成処理を行なう。すなわち、時刻t₁、t₂、t₃でそれぞれ送信された演奏データは、順次に受信されてRAM16内の受信バッファに書込まれる。RAM16内の参照スコアと称する記憶部には、受信バッファから演奏データの指示する実歌唱開始時刻順に演奏データが転送され、順次に書込まれる。例えば、演奏データS₁、S₂、S₃の順に書込まれる。そして、参照スコアの演奏データに基づいて実歌唱開始時刻順に歌唱合成スコアが形成される。例えば、演奏データS₁に基づいて歌唱合成スコアSC₁が形成され、演奏データS₂に基づいて歌唱合成スコアSC₂が形成される。この後、図9で述べたように歌唱合成スコアSC₁、SC₂…に従って歌唱合成が行なわれる。

30

【0063】

上記したのは、演奏データの送受信が実歌唱開始時刻順に行なわれた場合の参照スコア及び歌唱合成スコアの形成処理であるが、演奏データの送受信が実歌唱開始時刻順に行なわれなかつた場合の参照スコア及び歌唱合成スコアの形成処理は、図11、12に示すようになる。すなわち、演奏データS₁、S₃、S₄が時刻t₁、t₂、t₃にそれぞれ送信され、順次に受信されたものとすると、参照スコアには、図11に示すように演奏データS₁が書込まれた後、演奏データS₃、S₄が順次に書込まれ、演奏データS₁、S₃にそれぞれ基づいて歌唱合成スコアSC₁、SC_{3a}が形成される。参照スコアにおける2回目以降の演奏データの書き込みについて、図10、11に示したように単に追加的に書込むのを「追加」と称し、図12に示したように演奏データ間に挿入するように書込むのを「挿入」と称する。この後、時刻t₄に演奏データS₂が送信され、受信されたものとすると、図12に示すように参照スコアにおいて演奏データS₁及びS₃の間に演奏データS₂が挿入される。このとき、演奏データの挿入が起きた実歌唱開始時刻より後の歌唱合成スコアが破棄され、演奏データの挿入が起きた実歌唱開始時刻以降の演奏データに基づ

40

50

いて歌唱合成スコアが形成される。例えば、歌唱合成スコア $S C_{3,a}$ が破棄され、演奏データ S_2 、 S_3 にそれぞれ基づいて歌唱合成スコア $S C_2$ 、 $S C_{3,b}$ が形成される。

【0064】

図13は、ステップ42における演奏データに基づく歌唱合成スコアの形成例及びステップ44における歌唱音声の合成例を示すものである。歌唱合成スコア $S C$ は、RAM16内に形成されるもので、音韻トラック T_p と、遷移トラック T_R と、ビブラートトラック T_B とからなっている。歌唱合成スコア $S C$ については、演奏データを受信するたびにデータの更新や追加が行なわれる。

【0065】

一例として、演奏データ S_1 、 S_2 、 S_3 が図1(B)で示したと同様に「さ(s a) : $C_3 : T_1 \dots$ 」、「い(i) : $D_3 : T_2 \dots$ 」、「た(t a) : $E_3 : T_3 \dots$ 」を示すものとすると、音韻トラック T_p には、図13, 14に示すような情報が記憶される。すなわち、歌唱順に無音(Sil)、無音から子音「s」への遷移(Sil_s)、子音「s」から母音「a」への遷移(s_a)、母音(a)...等の情報が配置される。Silの情報は、開始時刻(Begin Time = T11)、継続時間(Duration = D11)及び音韻(PhU = Sil)を表わす情報からなる。Sil_sの情報は、開始時刻(Begin Time = T12)、継続時間(Duration = D12)先行音韻(PhU1 = Sil)及び後続音韻(PhU2 = s)を表わす情報からなる。s_aの情報は、開始時刻(Begin Time = T13)、継続時間(Duration = D13)、先行音韻(PhU1 = s)及び後続音韻(PhU2 = a)を表わす情報からなる。aの情報は、開始時刻(Begin Time = T14)、継続時間(Duration = D14)及び音韻(PhU = a)を表わす情報からなる。
10

【0066】

Sil_s , s_a 等の音韻遷移の継続時間を表わす継続時間情報は、先行音韻の時間長と後続音韻の時間長との組合せからなり、両時間長の境界を時間区切り情報として保持している。従って、この時間区切り情報を利用することにより音源回路28に対して先行音韻の継続時間と後続音韻の開始時刻及び継続時間とを指示可能である。例えば、 Sil_s の継続時間情報に基づいて無音の継続時間と子音「s」の歌唱開始時刻 $T_{1,1}$ 及び歌唱継続時間とを指示可能であり、 s_a の継続時間情報に基づいて子音「s」の継続時間と母音「a」の歌唱開始時刻 T_1 及び歌唱継続時間とを指示可能である。
20

【0067】

遷移トラック T_R には、図13, 15に示すような情報が記憶される。すなわち、遷移状態の進行順に遷移状態なし(NONEと表記する)、アタック遷移状態(Attack)、NONE、ノート遷移状態(NtN)、NONE、リリース遷移状態(Release)、NONE...等の状態情報が配置される。遷移トラック T_R の状態情報は、演奏データ及び音韻トラック T_p の情報に基づいて形成されるもので、Attackの状態情報は、音韻トラック T_p の「s」から「a」への音韻遷移の情報に対応し、NtNの状態情報は、音韻トラック T_p の「a」から「i」への音韻遷移の情報に対応し、Releaseの状態情報は、音韻トラック T_p の「a」から「Sil」への音韻遷移の情報に対応する。各状態情報は、対応する音韻遷移の情報に基づいて合成される歌唱音声に対して音高、振幅等の微小変化を付加するために用いられる。なお、図13の例では、音韻トラック T_p の「t」から「a」への音韻遷移に対応するNtNの状態情報を配置しないようにした。
30

【0068】

図15に示すように、1番目のNONEの状態情報は、開始時刻(Begin Time = T21)、継続時間(Duration = D21)及び遷移インデックス(Index = NONE)を表わす情報からなる。Attackの状態情報は、開始時刻(Begin Time = T22)、継続時間(Duration = D22)、遷移インデックス(Index = Attack)及び遷移インデックスのタイプ(例えば「Normal」、Type = Type22)を表わす情報からなる。2番目のNONEの状態情報は、開始時刻及び継続時間がそれぞれT23及びD23となる点を除き1番目のNONEの状態情報と同様である。NtNの状態情報は、開始時刻(Begin Time = T24)、継続時間(Duration = D24)、遷移インデックス(Index = NtN)及び遷移インデックスのタイプ(例えば「
40

Normal」、Type = Type 2 4) を表わす情報からなる。3番目のNONEの状態情報は、開始時刻及び継続時時間がそれぞれT 2 5 及びD 2 5 となる点を除き1番目のNONEの状態情報と同様である。Releaseの状態情報は、開始時刻 (Begin Time = T 2 6) 、継続時間 (Duration = D 2 6) 、遷移インデックス (Index = Release) 及び遷移インデックスのタイプ (例えば「Normal」、Type = Type 2 6) を表わす情報からなる。

【0069】

ビブラートトラック T_B には、図13, 16に示すような情報が記憶される。すなわち、ビブラートイベントの進行順にビブラートオフ (Vibrato OFF) 、ビブラートオン (Vibrato ON) 、ビブラートオフ (Vibrato OFF) ...等の情報が配置される。1番目のビブラートオフの情報は、開始時刻 (Begin Time = T 3 1) 、継続時間 (Duration = D 3 1) 10 及び遷移インデックス (Index = OFF) を表わす情報からなる。ビブラートオンの情報は、開始時刻 (Begin Time = T 3 2) 、継続時間 (Duration = D 3 2) 、遷移インデックス (Index = ON) 及びビブラートのタイプ (例えば「Normal」、Type = Type 3 2) を表わす情報からなる。2番目のビブラートオフの情報は、開始時刻及び継続時間がそれぞれT 3 3 及びD 3 3 となる点を除き1番目のビブラートオフ情報と同様である。

【0070】

ビブラートオンの情報は、音韻トラック T_p における「た」の母音「a」の情報に対応するもので、「a」の情報に基づいて合成される歌唱音声に音高、振幅等のビブラート的変化を付加するために用いられる。ビブラートオンの情報において、開始時刻を歌唱音声「a」の発生開始時刻 T_3 から遅延時間 D_L だけ遅く設定することにより遅延ビブラートを実現することができる。なお、図14～16に示した $T_{11} \sim T_{14}$, $T_{21} \sim T_{26}$, $T_{31} \sim T_{33}$ 等の開始時刻や $D_{11} \sim D_{14}$, $D_{21} \sim D_{26}$, $D_{31} \sim D_{33}$ 等の継続時間は、テンポクロック信号 T_{CL} のクロック数で適宜設定することができる。 20

【0071】

上記したような歌唱合成スコア SC と演奏データ $S_1 \sim S_3$ を用いると、ステップ4 4 の歌唱合成処理において、図13に示すような歌唱音声の合成が可能となる。音韻トラック T_p の Sil の情報に基づいて歌唱開始前の無音時間を実現した後、 T_p の Sil_s の情報と演奏データ S_1 中の C_3 の音高情報とに対応する音源制御情報を図6 (B) の音韻遷移 D_{B14b} から読み出して音源回路2 8 を制御することにより時刻 T_{11} で子音「s」の発生を開始させる。このときの制御期間は、 T_p の Sil_s の情報の指示する継続時間に対応する。次に、 T_p の s_a の情報と S_1 中の C_3 の音高情報とに対応する音源制御情報を D_{B14b} から読み出して音源回路2 8 を制御することにより時刻 T_1 で母音「a」の発生を開始させる。このときの制御期間は、 T_p の s_a の情報の指示する継続時間に対応する。この結果、歌唱音声 S_{S1} として「さ (s a)」が発生される。 30

【0072】

この後、 T_p の a の情報と S_1 の C_3 の音高情報とに対応する音源制御情報を図5の音韻 D_{B14a} から読み出して音源回路2 8 を制御することにより母音「a」の発生を継続する。このときの制御期間は、 T_p の a の情報の指示する継続時間に対応する。次に、 T_p の a_i の情報と S_2 中の D_3 の音高情報とに対応する音源制御情報を D_{B14b} から読み出して音源回路2 8 を制御することにより時刻 T_2 で母音「a」の発生を終了させると共に母音「i」の発生を開始させる。このときの制御期間は、 T_p の a_i の情報の指示する継続時間に対応する。 40

【0073】

この後、上記したと同様にして T_p の i 及び D_3 に対応する音源制御情報と T_p の i_t 及び D_3 に対応する音源制御情報を順次に読み出して音源回路2 8 を制御することにより母音「i」の発音を時刻 T_{31} まで継続し、時刻 T_{31} から子音「t」の発生を開始させる。そして、 T_p の t_a 及び E_3 に対応する音源制御情報に応じて音源回路2 8 を制御することにより時刻 T_3 から母音「a」の発生を開始させた後、 T_p の a 及び E_3 に対応する音源制御情報と、 T_p の a_Sil 及び E_3 に対応する音源制御情報を順次に読み出して音源回路2 8 を制御することにより母音「a」の発生を時刻 T_4 まで継続し、時刻 T_4 か 50

ら無音状態とする。この結果、歌唱音声 $S S_2, S S_3$ として、「い(i)」、「た(ta)」が順次に発生されることになる。

【0074】

上記のような歌唱音声の発生に伴って演奏データ $S_1 \sim S_3$ 中の情報と遷移トラック T_R の情報に基づいて歌唱音声制御が行なわれる。すなわち、時刻 T_1 の前後には、 T_R の Attack の状態情報と T_p の s_a の情報と S_1 中の C_3 の音高情報とに対応する音源制御情報を図 7 の状態遷移 $DB14c$ から読出して音源回路 28 を制御することにより歌唱音声「 s_a 」に対して音高、振幅等の微小変化を付加する。このときの制御期間は、Attack の状態情報の指示する継続時間に対応する。また、時刻 T_2 の前後には T_R の NtN の状態情報と T_p の a_i の情報と、 S_2 中の D_3 の音高情報とに対応する音源制御情報を $DB14c$ から読出して音源回路 28 を制御することにより歌唱音声「 a_i 」に対して音高、振幅等の微小変化を付加する。このときの制御期間は、NtN の状態情報の指示する継続時間に対応する。さらに、時刻 T_4 の直前には、 T_R の Release の状態情報と T_p の a の情報と S_3 中の E_3 の音高情報とに対応する音源制御情報を $DB14c$ から読出して音源回路 28 を制御することにより歌唱音声「 a 」に対して音高、振幅等の微小変化を付加する。このときの制御期間は、Release の状態情報の指示する継続時間に対応する。上記のような歌唱音声制御によれば、アタック感、ノート遷移感又はリリース感のある自然な歌唱合成が可能となる。10

【0075】

また、上記のような歌唱音声の発生に伴って演奏データ $S_1 \sim S_3$ 中の情報及びビブラートトラック T_B の情報に基づいて歌唱音声制御が行われる。すなわち、時刻 T_3 から遅延時間 DL だけ遅れた時刻になると、 T_B のビブラートオンの情報と T_p の a の情報と S_3 中の E_3 の音高情報とに対応した音源制御情報を図 8 のビブラート $DB14d$ から読出して音源回路 28 を制御することにより歌唱音声「 a 」に対して音高、振幅等のビブラート的变化を付加し、このような付加を時刻 T_4 まで続ける。このときの制御期間は、ビブラートオンの情報の指示する継続時間に対応する。また、ビブラートの深さや速さは、 S_3 中のビブラートタイプの情報に応じて決定される。上記のような歌唱音声制御によれば、所望の歌唱個所にビブラートを付加して自然な歌唱合成を行なうことができる。20

【0076】

次に、図 17 を参照して演奏データ受信・歌唱合成スコア形成処理を説明する。ステップ 50 では、初期化処理を行ない、例えば RAM16 内の受信回数カウンタ n に 0 をセットする。30

【0077】

ステップ 52 では、カウンタ n の値 n を 1 増加 ($n + 1$) する。そして、ステップ 54 で変数 m をカウンタ n の値 n とし、 $m = n$ 番目の演奏データを受信して RAM16 内の受信バッファに書込む。

【0078】

ステップ 56 では、 $m = n$ 番目の演奏データがデータエンドか判定する。ステップ 54 で $m = 1$ 番目の演奏データを受信したときは、ステップ 56 の判定結果が否定的 (N) となり、ステップ 58 に移る。ステップ 58 では、 $m = n$ 番目の演奏データを受信バッファから読出して RAM16 内の参照スコアに書込む。参照スコアに $m = 1$ 番目の演奏データを書込んだ後は、図 10 ~ 12 で前述したように参照スコアに演奏データを「追加」又は「挿入」の形で書込むことになる。40

【0079】

次に、ステップ 60 では、 $n > 1$ か判定する。 $m = 1$ 番目の演奏データを受信したときは、ステップ 60 の判定結果が否定的 (N) となり、ステップ 52 に戻る。ステップ 52 では、 $n = 2$ とし、ステップ 54 では、 $m = 2$ 番目の演奏データを受信し、受信バッファに書込む。そして、ステップ 56 を経てステップ 58 に来ると、 $m = 2$ 番目の演奏データを参照スコアに追加する。

【0080】

次に、ステップ 6 0 で $n > 1$ か判定すると、 $n = 2$ であるので、判定結果が肯定的 (Y) となり、ステップ 6 1 の歌唱合成スコア形成処理を行なう。ステップ 6 1 の処理については、図 1 8 を参照して後述するが、概略的に説明すると、次のようになる。すなわち、ステップ 6 2 では、 $m = n - 1$ 番目の演奏データを参照スコアに挿入したか判定する。例えば、 $m = 1$ 番目の演奏データについては、挿入ではないので、ステップ 6 2 の判定結果が否定的 (N) となり、ステップ 6 4 に移る。ステップ 6 4 では、 $m = n - 1$ 番目の演奏データについて歌唱合成スコアを形成する。例えば、ステップ 5 4 で $m = 2$ 番目の演奏データを受信したときは、ステップ 6 4 で $m = 1$ 番目の演奏データについて歌唱合成スコアが形成される。

【0081】

ステップ 6 4 の処理が終わると、ステップ 5 2 に戻り、上記したと同様にして演奏データの受信及び参照スコアへの演奏データの書き込みを行なう。例えば、ステップ 6 4 で $m = 1$ 番目の演奏データについて歌唱合成スコアを形成した後は、ステップ 5 4 で $m = 3$ 番目の演奏データを受信し、ステップ 5 8 で $m = 3$ 番目の演奏データを参照スコアに追加又は挿入する。

【0082】

ステップ 6 2 の判定結果が肯定的 (Y) であったときは、参照スコアに $m = n - 1$ 番目の演奏データを挿入した場合であり、ステップ 6 6 に移る。ステップ 6 6 では $m = n - 1$ 番目の演奏データより実歌唱開始時刻が後の歌唱合成スコアを破棄し、 $m = n - 1$ 番目の演奏データ以降の演奏データについて歌唱合成スコアを形成する。例えば、図 1 1, 1 2 に示すように演奏データ S_1, S_3, S_4 を受信した後、演奏データ S_2 を受信すると、ステップ 5 8 では、 $m = 4$ 番目の演奏データ S_2 を参照スコアに挿入する。そして、ステップ 6 0 を経てステップ 6 2 に来ると、 $m = 4 - 1 = 3$ 番目の演奏データ S_4 は、参照スコアに追加されたものであるので、ステップ 6 2 の判定結果が否定的 (N) となり、ステップ 6 4 を経てステップ 5 2 に戻る。そして、ステップ 5 4 で $m = 5$ 番目の演奏データを受信した後、ステップ 5 6, 5 8, 6 0 を経てステップ 6 2 に来ると、 $m = 4$ 番目の演奏データ S_2 は、参照スコアに挿入されたものであるので、ステップ 6 2 の判定結果が肯定的 (Y) となり、ステップ 6 6 において、 $m = 4$ 番目の演奏データ S_2 より実歌唱開始時刻が後の歌唱合成スコア (図 1 2 の $SC_{3,a}$ 等) を破棄し、 $m = 4$ 番目の演奏データ S_2 以降の演奏データ (図 1 2 の S_2, S_3, S_4) について歌唱合成スコアを形成する。

【0083】

ステップ 6 6 の処理が終わると、ステップ 5 2 に戻り、上記したと同様の処理を繰返す。 $m = n$ 番目の演奏データがデータエンドになると、ステップ 5 6 の判定結果が肯定的 (Y) となり、ステップ 6 8 で終端処理 (例えばエンド情報の付加等) を行なう。ステップ 6 8 の後は、図 3 のステップ 4 4 の歌唱合成処理を行なう。

【0084】

図 1 8 は、歌唱合成スコア形成処理を示すものである。ステップ 7 0 では、参照スコアから図 4 に示したような演奏情報を含む演奏データを取得する。ステップ 7 2 では、取得した演奏データに含まれる演奏情報を分析する。ステップ 7 4 では、分析された演奏情報及び保存された管理データ (先行演奏データの管理データ) に基づいて歌唱合成スコア形成のための管理データを作成する。ステップ 7 4 の処理については、図 1 9 を参照して後述する。

【0085】

次に、ステップ 7 6 では、取得に係る演奏データが参照ストアに書き込まれた際に挿入されたか判定する。この判定の結果が肯定的 (Y) であれば、ステップ 7 8 において、取得に係る演奏データより実歌唱開始時刻が後の歌唱合成スコアを破棄する。

【0086】

ステップ 7 8 の処理が終わったとき又はステップ 7 6 の判定結果が否定的 (N) であったときは、ステップ 8 0 で音韻トラック形成処理を行なう。ステップ 8 0 の処理では、演奏情報と、ステップ 7 4 で作成された管理データと、保存されたスコアデータ (先行演奏

10

20

30

40

50

データのスコアデータ)とに基づいて音韻トラック T_p を形成するが、このような処理の詳細については、図22を参照して後述する。

【0087】

ステップ82では、演奏情報と、ステップ74で作成された管理データと、保存されたスコアデータと音韻トラック T_p とに基づいて遷移トラック T_R を形成する。ステップ82の処理の詳細については、図34を参照して後述する。

【0088】

ステップ84では、演奏情報と、ステップ74で作成された管理データと、保存されたスコアデータと音韻トラック T_p とに基づいてビブラートトラック T_B を形成する。ステップ84の処理の詳細については、図37を参照して後述する。

10

【0089】

ステップ86では、演奏情報と、ステップ74で作成された管理データと、音韻トラック T_p と、遷移トラック T_R と、ビブラートトラック T_B とに基づいて次の演奏データのためのスコアデータを形成し、保存する。スコアデータには、先行母音からのNtN遷移時間長が含まれる。このNtN遷移時間長は、図36に示すように先行ノート(先行母音)にかかる時間長 T_1 と後続ノート(本演奏データ)にかかる時間長 T_2 との組合せからなるもので、両時間長の境界を時間区切り情報として保持している。NtN遷移時間長を算出するには、演奏情報中の歌唱音韻と歌唱音高と歌唱ノート遷移タイプ(例えば「Normal」)とに対応するNtNの状態遷移時間長を図7の状態遷移DB14cから読み出し、この状態遷移時間長に対して演奏情報中の歌唱ノート遷移伸縮率を乗算すればよい。この乗算の結果として得られるNtN遷移時間長は、図13, 15に示したNtNの状態情報中の継続時間情報として用いられる。

20

【0090】

図19は、管理データ作成処理を示すものである。管理データとしては、図20, 21に示すように音韻状態(PhU State)、音素(Phoneme)、音高(Pitch)、現ノートオン(Current Note On)、現ノートデュレーション(Current Note Duration)、現ノートオフ(Current Note Off)、フルデュレーション(Full Duration)、イベント状態(Event State)等の情報がある。

【0091】

ステップ90で演奏データを取得した後、ステップ92では、演奏データ中の歌唱音韻を分析する。音韻状態情報は、子音と母音との組合せ、母音のみ又は有声子音のみのいずれかを表わす。以下では、便宜上、子音と母音との組合せの場合をPhU State = Consonant Vowel とし、母音のみ又は有声子音のみの場合をPhU State = Vowel とする。音素情報は、音素名(子音名及び/又は母音名)、子音のカテゴリ(鼻音、破裂音、半母音、その他)、子音の有声又は無声等を表わす。

30

【0092】

ステップ94では、演奏データ中の歌唱音高を分析し、この歌唱音高を音高情報とする。ステップ96では、演奏データ中の実歌唱時刻を分析し、実歌唱開始時刻を現ノートオン情報とする。また、実歌唱長を現ノートデュレーション情報とし、実歌唱開始時刻から実歌唱長だけ経過した時刻を現ノートオフ情報とする。

40

【0093】

現ノートオン情報としては、実歌唱開始時刻を修正した時刻を採用してもよい。例えば、乱数発生処理等により実歌唱開始時刻(実線で示す)の前後に及ぶ所定の時間範囲(2本の破線で示す)内で t だけランダムに変更した時刻(実歌唱開始時刻を t_0 とすると、 $t_0 \pm t$ の時刻)を現ノートオン情報としてもよい。

【0094】

ステップ98では、先行演奏データの管理データを用いて先行歌唱イベントとの関係で本演奏データの歌唱時刻を分析する。先行演奏データの管理データにおいて、先行イベントナンバ(Event Number)情報は、先行演奏データの並べ替えを済ませた受信ナンバを表わし、先行スコアデータ(Score Data)は、先行演奏データについて歌唱合成スコアを

50

形成した際にステップ 8 6 で形成し、保存されたスコアデータであり、先行ノートオフ (Note Off) 情報は、先行実歌唱終了時刻を表わす。イベントステート情報は、先行ノートオフ情報及び現ノートオン情報に基づいて決定された先行歌唱イベントと現歌唱イベントの接続状況（無音の有無）を表わす。以下では、便宜上、図 2 0 に示すように先行歌唱イベントと現歌唱イベントとが連続する場合（無音なしの場合）を Event State = Transition とし、図 2 1 に示すように先行歌唱イベントと現歌唱イベントとの間に無音が入る場合を Event State = Attack とする。フルデュレーション情報は、先行ノートオフ情報の指示する先行実歌唱終了時刻から現ノートオフ情報の指示する実歌唱終了時刻までの時間長を表わす。

【0095】

10

次に、図 2 2 を参照して音韻トラック形成処理を説明する。ステップ 1 0 0 では、演奏情報（演奏データの内容）、管理データ及びスコアデータを取得する。ステップ 1 0 2 では、取得データに基づいて図 6 (B) の音韻遷移 DB14b から音韻遷移時間長を獲得する（読み出す）。ステップ 1 0 2 の処理については、図 2 3 を参照して後述する。

【0096】

ステップ 1 0 4 では、管理データに基づいて Event State = Attack か判定する。この判定の結果が肯定的 (Y) であれば、前に無音 (Silence) が存在することになり、ステップ 1 0 6 で Silence 歌唱長を算出する。ステップ 1 0 6 の処理については、図 2 4 を参照して後述する。

【0097】

20

ステップ 1 0 4 の判定結果が否定的 (N) であったときは、Event State = Transition で前に母音が存在することになり、ステップ 1 0 8 で先行 Vowel (母音) 歌唱長を算出する。ステップ 1 0 8 の処理については、図 2 8 を参照して後述する。

【0098】

ステップ 1 0 6 又は 1 0 8 の処理が終わったときは、ステップ 1 1 0 で Vowel 歌唱長を算出する。ステップ 1 1 0 の処理については、図 3 2 を参照して後述する。

【0099】

図 2 3 は、ステップ 1 0 2 の音韻遷移時間長獲得処理を示すもので、ステップ 1 1 2 では、管理データ及びスコアデータを取得する。そして、ステップ 1 1 4 では、すべての音韻遷移時間長（後述のステップ 1 1 6, 1 2 2, 1 2 4, 1 2 6, 1 3 0, 1 3 2, 1 3 4 30 で獲得される音韻遷移時間長）を初期化する。

【0100】

ステップ 1 1 6 では、管理データに基づいて V_Sil (母音から無音へ) の音韻遷移時間長を DB14b から獲得する。一例として、母音が「a」であり、その音高が「P1」であるとすると、DB14b からは、「a_Sil」と「P1」とに対応した音韻遷移時間長が獲得される。ステップ 1 1 6 の処理は、日本語が母音で終わることと関連している。

【0101】

ステップ 1 1 8 では、管理データに基づいて Event State = Attack か判定する。この判定の結果が肯定的 (Y) であれば、ステップ 1 2 0 で管理データに基づいて PhU State = Consonant Vowel か判定する。この判定の結果が肯定的 (Y) であれば、ステップ 1 2 2 で管理データに基づいて DB14b から Sil_C (無音から子音へ) の音韻遷移時間長を獲得する。この後、ステップ 1 2 4 では、管理データに基づいて DB14b から C_V (子音から母音へ) の音韻遷移時間長を獲得する。

40

【0102】

ステップ 1 2 0 の判定結果が否定的 (N) であったときは、PhU State = Vowel であったことになり、ステップ 1 2 6 で管理データに基づいて DB14b から Sil_V (無音から母音へ) の音韻遷移時間長を獲得する。なお、ステップ 1 2 2 ~ 1 2 6 における音韻遷移時間長の具体的な獲得法は、ステップ 1 1 6 で述べたと同様である。

【0103】

ステップ 1 1 8 の判定結果が否定的 (N) であったときは、ステップ 1 2 0 と同様にして

50

ステップ128でPhU State = Consonant Vowelか判定する。この判定の結果が肯定的(Y)であれば、ステップ130で管理データ及びスコアデータに基づいてpV_C(先行母音から子音へ)の音韻遷移時間長をDB14bから獲得する。一例として、スコアデータにより先行母音が「a」であり、管理データにより子音が「s」で、その音高が「P2」であるとすると、DB14bからは、「a_s」とび「P2」に対応した音韻遷移時間長が獲得される。この後、ステップ132では、ステップ116で述べたと同様にして管理データに基づいてDB14bからC_V(子音から母音へ)の音韻遷移時間長を獲得する。

【0104】

ステップ128の判定結果が否定的(N)であったときは、ステップ134に移る。ステップ134では、ステップ130で述べたと同様にして管理データ及びスコアデータに基づいてDB14bからpV_V(先行母音から母音へ)の音韻遷移時間長を獲得する。

【0105】

図24は、ステップ106のSilence歌唱長算出処理を示すもので、ステップ136では、演奏情報、管理データ及びスコアデータを獲得する。

【0106】

ステップ138では、PhU State = Consonant Vowelか判定する。この判定の結果が肯定的(Y)であれば、ステップ140でConsonant歌唱長を算出する。この場合、図25に示すように、子音歌唱時間は、無音から子音への音韻遷移時間長内の子音部と、Consonant歌唱長と、子音から母音への音韻遷移時間長内の子音部との加算によって決まる。従って、Consonant歌唱長は、子音歌唱時間の一部となる。

【0107】

図25には、演奏情報に含まれる歌唱子音伸縮率が1より大きい場合においてConsonant歌唱長を決定する例を示す。この場合、Sil_Cの子音長とC_Vの子音長とを加算したもの的基本単位とし、これに歌唱子音伸縮率を乗算したものをConsonant歌唱長Cとする。そして、Sil_CとC_Vとの間にConsonant歌唱長Cを介在させることによって子音歌唱時間を伸長する。

【0108】

図26には、歌唱子音伸縮率が1より小さい場合においてConsonant歌唱長を決定する例を示す。この場合、Sil_Cの子音長とC_Vの子音長とにそれぞれ歌唱子音伸縮率を乗算して各々子音長を短縮する。この結果、Sil_Cの子音長とC_Vの子音長とからなる子音歌唱時間が短縮される。

【0109】

次に、ステップ142では、Silence歌唱長を算出する。図27に示すように、無音時間は、先行母音から無音への音韻遷移時間長の無音部と、Silence歌唱長と、無音から子音への音韻遷移時間長の無音部と、子音歌唱時間との加算、あるいは先行母音から無音への音韻遷移時間長の無音部と、Silence歌唱長と、無音から母音への音韻遷移時間長の無音部との加算によって決まる。従って、Silence歌唱長は、無音時間の一部となる。ステップ142では、歌唱の順に従い、C_V内の子音部と母音部との境界又はSil_V内の無音部と母音部との境界が実歌唱開始時刻(Current Note On)となるようにSilence歌唱長を算出する。すなわち、本演奏データの母音の歌唱開始時刻が実歌唱開始時刻と一致するようにSilence歌唱長を算出する。

【0110】

図27(A)～(C)には、互いに異なる音韻接続パターンを示す。(A)のパターンは、例えば先行母音「あ」-無音-「さ(sa)」に対応し、子音「s」を伸長するためにConsonant歌唱長Cが挿入されている。(B)のパターンは、例えば先行母音「あ」-無音-「ぱ(pa)」に対応する。(C)のパターンは、例えば先行母音「あ」-無音-「い(i)」に対応する。

【0111】

図28は、ステップ108の先行Vowel歌唱長算出処理を示すもので、ステップ146で

10

20

30

40

50

は、演奏情報、管理データ及びスコアデータを取得する。

【0112】

ステップ148では、PhU State = Consonant Vowel か判定する。この判定の結果が肯定的(Y)であれば、ステップ150でConsonant歌唱長を算出する。この場合、図29に示すように、子音歌唱時間は、先行母音から子音への音韻遷移時間長内の子音部と、Consonant歌唱長と、子音から母音への音韻遷移時間長内の子音部との加算によって決まる。従って、Consonant歌唱長は、子音歌唱時間の一部となる。

【0113】

図29には、演奏情報に含まれる歌唱子音伸縮率が1より大きい場合においてConsonant歌唱長を決定する例を示す。この場合、pV_Cの子音長とC_Vの子音長とを加算したもの 10 を基本単位とし、これに歌唱子音伸縮率を乗算したものをConsonant歌唱長Cとする。そして、pV_CとC_Vとの間にConsonant歌唱長Cを介在させることによって子音歌唱時間を伸長する。

【0114】

図30には、歌唱伸縮率が1より小さい場合においてConsonant歌唱長を決定する例を示す。この場合、pV_Cの子音長とC_Vの子音長とにそれぞれ歌唱伸縮率を乗算して各々の子音長を短縮する。この結果、pV_Cの子音長とC_Vの子音長とからなる子音歌唱時間が短縮される。

【0115】

次に、ステップ152では、先行Vowel歌唱長を算出する。図31に示すように、先行母音歌唱時間は、X(Sil、子音又は母音)から先行母音への音韻遷移時間長の母音部と、先行Vowel歌唱長と、先行母音から子音又は母音への音韻遷移時間長の母音部との加算によって決まる。従って、先行Vowel歌唱長は、先行母音歌唱時間の一部となる。また、本演奏データを受信したことにより先行演奏データと本演奏データとの接続が明確になったので、先行演奏データに基づいて形成されたVowel歌唱長及びV_Silは破棄する。すなわち、後述する図32のVowel歌唱長算出処理で用いられる「次の演奏データとの間に無音が挿入される」という仮定は破棄される。ステップ152では、歌唱の順に従い、C_V内の子音部と母音部との境界又はpV_V内の先行母音部と母音部との境界が実歌唱開始時刻(Current Note On)となるように先行Vowel歌唱長を算出する。すなわち、本演奏データの母音の歌唱開始時刻が実歌唱開始時刻と一致するように先行Vowel歌唱長を算出する。 30

【0116】

図31(A)～(C)には、互いに異なる音韻接続パターンを示す。(A)のパターンは、例えば先行母音「あ」-「さ(sa)」に対応し、子音「s」を伸長するためにConsonant歌唱長Cが挿入されている。(B)のパターンは、例えば、先行母音「あ」-「ぱ(pa)」に対応する。(C)のパターンは、例えば、先行母音「あ」-「い(i)」に対応する。

【0117】

図32は、ステップ110のVowel歌唱長算出処理を示すもので、ステップ154では、演奏情報、管理データ及びスコアデータを取得する。

【0118】

ステップ156では、Vowel歌唱長を算出する。この場合、次の演奏データを受信しないと、母音の接続部が明確にならないため、「次の演奏データとの間に無音が挿入される」と仮定し、図33に示すように母音部にV_Silを接続してVowel歌唱長を算出する。この時点において、母音歌唱時間は、Xから母音への音韻遷移時間長の母音部と、Vowel歌唱長と、母音から無音への音韻遷移時間長の母音部との加算によって一時的に決まる。従って、Vowel歌唱長は、母音歌唱時間の一部となる。ステップ156では、歌唱の順に従い、V_Sil内の母音部と無音部との境界が実歌唱終了時刻(Current Note Off)となるようにVowel歌唱長を決定する。

【0119】

次の演奏データを受信したときは、本演奏データとの接続状態(Event State)が明確に 50

なり、次の演奏データのEvent State = Attackの場合には本演奏データのVowel歌唱長が更新されず、次の演奏データのEvent State = Transitionの場合には前述のステップ152の処理によって本演奏データのVowel歌唱長が更新される。

【0120】

図34は、遷移トラック形成処理を示すもので、ステップ160では、演奏情報、管理データ、スコアデータ及び音韻トラックのデータを取得する。

【0121】

ステップ162では、Attack遷移時間長を算出する。このためには、演奏情報及び管理データに基づいて歌唱アタックタイプと音韻と音高とに対応したAttack状態の状態遷移時間長を図7の状態遷移DB14cから獲得する。そして、獲得した状態遷移時間長に対して演奏情報中の歌唱アタック伸縮率を乗算したものをAttack遷移時間長（アタック部の継続時間）とする。
10

【0122】

ステップ164では、Release遷移時間長を算出する。このためには、演奏情報及び管理データに基づいて歌唱リリースタイプと音韻と音高とに対応したRelease状態の状態遷移時間長をDB14cから獲得する。そして、獲得した状態遷移時間長に対して演奏情報中の歌唱リリース伸縮率を乗算したものをRelease遷移時間長（リリース部の継続時間）とする。

【0123】

ステップ166では、NtN遷移時間長を獲得する。すなわち、図18のステップ86で保存されたスコアデータ中から、先行母音からのNtN遷移時間長（ノート遷移部の継続時間）を獲得する。
20

【0124】

ステップ168では、Event State = Attackか判定する。この判定の結果が肯定的（Y）であれば、ステップ170で無音部に対応するNONE遷移時間長（「NONEn遷移時間長」と称する）を算出する。すなわち、PhU State = Consonant Vowel の場合は、図35（A）、（B）に示すように子音の歌唱開始時刻とAttack遷移開始時刻（Attack遷移時間長の先端位置）とが一致するようにNONEn遷移時間長を算出する。図35において、（A）の例が（B）の例と異なるのは、子音歌唱時間にConsonant歌唱長Cを挿入した点である。PhU State = Vowelの場合は、図35（C）に示すように母音の歌唱開始時刻とAttack遷移開始時刻とが一致するようにNONEn遷移時間長を算出する。
30

【0125】

ステップ172では、定常部に対応するNONE遷移時間長（「NONEs遷移時間長」と称する）を算出する。この場合、次の演奏データを受信しないと、NONEs遷移時間長に続く接続状態が明確にならないため、「次の演奏データとの間に無音が挿入される」と仮定し、図35に示すようにRelease遷移を接続した状態でNONEs遷移時間長を算出する。すなわち、Release遷移終了時刻（Release遷移時間長の終端位置）がV_Silの終了時刻と一致するように先行演奏データの終了時刻とV_Silの終了時刻とAttack遷移時間長とRelease時間長とNONEn遷移時間長とに基づいてNONEs遷移時間長を算出する。

【0126】

ステップ168の判定結果が否定的（N）であったときは、ステップ174で先行演奏データの定常部に対応するNONE遷移時間長（「pNONEs遷移時間長」と称する）を算出する。本演奏データを受信したことにより先行演奏データとの接続状態が明確になったので、先行演奏データに基づいて形成されたNONEs遷移時間長及び先行Release遷移時間長は破棄する。すなわち、後述するステップ176の処理で用いられる「次の演奏データとの間に無音が挿入される」という仮定は破棄される。ステップ174では、図36（A）～（C）に示すようにPhU State = Consonant Vowel又はPhU State = Vowelのいずれの場合にも、先行母音からのNtN遷移時間長のT₁及びT₂の境界と本演奏データの実歌唱開始（Current Note On）とが一致するように本演奏データの実歌唱開始時刻及び実歌唱終了時刻とNtN遷移時間長とに基づいてpNONEs遷移時間長を算出する。図36におい
40
50

て、(A)の例が(B)の例と異なるのは、子音歌唱時間にConsonant歌唱長Cを挿入した点である。

【0127】

ステップ176では、定常部に対応するNONE遷移時間長(NONEs遷移時間長)を算出する。この場合、次の演奏データを受信しないと、NONEs遷移時間長に続く接続状態が明確にならないため、「次の演奏データとの間に無音が挿入される」と仮定し、図36に示すようにRelease遷移を接続した状態でNONEs遷移時間長を算出する。すなわち、先行母音からのNtN遷移時間長のT₁及びT₂の境界と本演奏データの実歌唱開始時刻(Current Note On)とが一致し且つRelease遷移終了時刻(Release遷移時間長の終端位置)とV_Silの終了時刻とが一致するように本演奏データの実歌唱開始時刻とV_Silの終了時刻と先行母音からのNtN遷移時間長とRelease遷移時間長とに基づいてNONEs遷移時間長を算出する。10

【0128】

図37は、ビブラートトラック形成処理を示すもので、ステップ180では、演奏情報、管理データ、スコアデータ及び音韻トラックのデータを取得する。

【0129】

ステップ182では、取得データに基づいてビブラートイベント継続か判定する。本演奏データの実歌唱開始時刻からビブラートが開始され、先行演奏データからビブラート付加状態が継続する場合は、ステップ182の判定結果が肯定的(Y)となり、ステップ184に移る。また、本演奏データの実歌唱開始時刻からビブラートが開始されるが、先行演奏データからビブラート付加状態が継続しない場合、あるいは本演奏データからビブラートが開始されない場合は、ステップ182の判定結果が否定的(N)となり、ステップ188に移る。20

【0130】

ビブラートは、複数の演奏データ(ノート)にまたがって歌唱されることが多い。本演奏データの実歌唱開始時刻からビブラートが開始されていても、図38(A)に示すように先行ノートからビブラート付加状態が継続する場合と、図38(D)、(E)に示すように現ノートの実歌唱開始時刻からビブラートが追加的に開始される場合とがありうる。同様に、無ビブラート(ビブラート非付加状態)についても、本演奏データの実歌唱開始時刻から無ビブラートが開始されている場合と、図38(B)に示すように先行ノートから無ビブラートが継続する場合と、図38(C)に示すように現ノートの実歌唱開始時刻から無ビブラートが開始される場合とがありうる。30

【0131】

ステップ188では、取得データに基づいて無ビブラートイベント継続か判定する。先行ノートから無ビブラート状態が継続する図38(B)のような場合は、ステップ188の判定結果が肯定的(Y)となり、ステップ190に移る。また、現ノートの実歌唱開始時刻から無ビブラートが開始されているが、先行ノートから無ビブラート状態が継続しない図38(C)のような場合、あるいは現ノートの実歌唱開始時刻から無ビブラートが開始されない場合は、ステップ188の判定結果が否定的(N)となり、ステップ194に移る。40

【0132】

ビブラートイベント継続の場合、ステップ184では、先行ビブラート時間長を破棄する。そして、ステップ186で先行ビブラート時間長と現ノートの実歌唱開始時刻から開始されるビブラートのビブラート時間長とを接続(加算)して新たなビブラート時間長を算出する。そして、ステップ194に移る。

【0133】

無ビブラートイベント継続の場合、ステップ190では、先行無ビブラート時間長を破棄する。そして、ステップ192で先行無ビブラート時間長と現ノートの実歌唱開始時刻から開始される無ビブラートの無ビブラート時間長とを接続(加算)して新たな無ビブラート時間長を算出する。そして、ステップ194に移る。50

【0134】

ステップ194では、ビブラート時間長の追加ありか判定する。この判定の結果が肯定的(Y)であれば、ステップ196で追加無ビブラート時間長を算出する。すなわち、ステップ186で算出されたビブラート時間長の終端から追加するビブラート時間長までの無ビブラート時間長を追加無ビブラート時間長として算出する。

【0135】

ステップ198では、追加ビブラート時間長を算出する。そして、ステップ194に戻り、それ以降の処理を上記したと同様に繰返す。この結果、複数の追加ビブラート時間長を算出可能となる。

【0136】

ステップ194の判定結果が否定的(N)であったときは、ステップ200で無ビブラート時間長を算出する。すなわち、実歌唱時間長(Current Note OnからCurrent Note Offまでの時間長)内で最終のビブラートイベントの最終時刻からV_Silの終了時刻までを無ビブラート時間長として算出する。

【0137】

上記したステップ142又は152では、本演奏データの母音の歌唱開始時刻が実歌唱開始時刻と一致するようにSilence歌唱長又は先行Vowel歌唱長を算出したが、より自然な歌唱合成を行なうこと目的として、次の(1)~(11)で述べるようにしてSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出してもよい。

【0138】

(1) 子音のカテゴリ(無声/有声破裂音、無声/有声摩擦音、鼻音、半母音等)毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。図39には、子音のカテゴリが鼻音又は半母音である場合に他の場合とはSilence歌唱長の決定の仕方を変えた例を示す。

【0139】

図39において、(A)の例は、音韻接続パターンとしては、先行母音「あ」-無音-「さ(s a)」に対応したもので、子音及び母音からなる音韻について子音(この例では「s」)を伸ばすためにConsonant歌唱長を挿入した状態でSilence歌唱長を算出した例である。(B)の例は、音韻接続パターンとしては、先行音韻「あ」-無音-「ぱ(p a)」に対応したもので、子音及び母音からなる音韻についてConsonant歌唱長を挿入せずにSilence歌唱長を算出した例である。(C)の例は、音韻接続パターンとしては、先行母音「あ」-無音-「な(n a)」に対応したもので、子音(鼻音又は半母音)及び母音からなる音韻について子音(この例では「n」)を伸ばすためにConsonant歌唱長Cを挿入した状態でSilence歌唱長を算出した例である。(D)の例は、Consonant歌唱長を挿入しない点を除き(C)と同様の例である。(E)の例は、音韻接続パターンとしては、先行音韻「あ」-無音-「い(i)」に対応したもので、母音のみからなる音韻についてSilence歌唱長を算出した(子音(鼻音)のみからなる音韻についても同様とする)例である。

【0140】

(A)、(B)及び(E)の例は、本演奏データの母音の歌唱開始時刻が実歌唱開始時刻と一致するようにSilence歌唱長を算出した例である。(C)及び(D)の例は、本演奏データの子音の歌唱開始時刻が実歌唱開始時刻と一致するようにSilence歌唱長を算出した例である。

【0141】

(2) 子音(「p」、「b」、「s」、「z」、「n」、「w」等)毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0142】

(3) 母音(「a」、「i」、「u」、「e」、「o」等)毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0143】

(4) 子音のカテゴリ(無声/有声破裂音、無声/有声摩擦音、鼻音、半母音等)毎に且

10

20

30

40

50

つ子音に続く母音（「a」、「i」、「u」、「e」、「o」等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、子音のカテゴリと母音との組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0144】

(5) 子音（「p」、「b」、「s」、「z」、「n」、「w」等）毎に且つ子音に続く母音毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、子音と母音との組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0145】

(6) 先行母音（「a」、「i」、「u」、「e」、「o」等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。 10

【0146】

(7) 先行母音（「a」、「i」、「u」、「e」、「o」等）毎に且つ先行母音に続く子音のカテゴリ（無声／有声破裂音、無声／有声摩擦音、鼻音、半母音等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、先行母音と子音のカテゴリとの組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0147】

(8) 先行母音（「a」、「i」、「u」、「e」、「o」等）毎に且つ先行母音に続く子音（「p」、「b」、「s」、「z」、「n」、「w」等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、先行母音と子音との組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。 20

【0148】

(9) 先行母音（「a」、「i」、「u」、「e」、「o」等）毎に且つ先行母音に続く母音（「a」、「i」、「u」、「e」、「o」等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、先行母音と母音との組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0149】

(10) 先行母音（「a」、「i」、「u」、「e」、「o」等）毎、先行母音に続く子音のカテゴリ（無声／有声破裂音、無声／有声摩擦音、鼻音、半母音等）毎及び子音に続く母音（「a」、「i」、「u」、「e」、「o」等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、先行母音と子音のカテゴリと母音との組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。 30

【0150】

(11) 先行母音（「a」、「i」、「u」、「e」、「o」等）毎、先行母音に続く子音（「p」、「b」、「s」、「z」、「n」、「w」等）毎及び子音に続く母音（「a」、「i」、「u」、「e」、「o」等）毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。すなわち、先行母音と子音と母音との組合せ毎にSilence歌唱長、先行Vowel歌唱長及びVowel歌唱長を算出する。

【0151】

この発明は、上記した実施形態に限定されるものではなく、種々の改変形態で実施可能なものである。例えば、次のような変更が可能である。 40

【0152】

(1) 上記した実施形態では、歌唱合成スコアの形成が完了した後、歌唱合成スコアに従って歌唱音声を合成するようにしたが、歌唱合成スコアを形成しつつ形成済みの歌唱合成スコアに従って歌唱音声を合成するようにしてもよい。このためには、例えば、演奏データの受信を割り込み処理により優先的に行ないつつ受信済みの演奏データに基づいて歌唱合成スコアを形成すればよい。

【0153】

(2) 上記した実施形態では、音源方式としてフォルマント合成方式を用いたが、波形処理方式等の他の方式を用いてもよい。

【0154】

(3) 上記した実施形態では、歌唱合成スコアを音韻トラック、遷移トラック及びビブラートトラックの3トラックで構成したが、1トラック構成としてもよい。このためには、例えば、音韻トラックに遷移トラック及びビブラートトラックの情報を適宜挿入すればよい。

【0155】

【発明の効果】

以上のように、この発明によれば、子音から母音に遷移する音韻について子音の立上りを母音の立上りより十分に早くすることができるので、人歌唱に近似した自然な歌唱合成が可能となる効果が得られる。

【0156】

また、音韻遷移時間長を用いて歌唱開始時刻や歌唱継続時間を決定するようにしたので、遷移する音韻について自然な歌唱合成が可能となる効果も得られる。

10

【図面の簡単な説明】

【図1】 人歌唱と、この発明の歌唱合成とで歌唱開始時刻を対比して示す図である。

【図2】 この発明の一実施形態に係る歌唱合成装置の回路構成を示すブロック図である。
。

【図3】 実施形態に係る歌唱合成処理を概略的に説明するためのフローチャートである。
。

【図4】 演奏データ内的情報を示す図である。

【図5】 音韻DB内的情報を示す図である。

20

【図6】 音韻遷移DB内的情報を示す図である。

【図7】 状態遷移DB内的情報を示す図である。

【図8】 ビブラートDB内的情報を示す図である。

【図9】 演奏データに基づく歌唱合成を説明するための図である。

【図10】 参照スコア及び歌唱合成スコアの形成状況を示す図である。

【図11】 参照スコアに演奏データを追加した場合の歌唱合成スコアの形成状況を示す図である。

【図12】 参照スコアに演奏データを挿入した場合の歌唱合成スコアの形成状況を示す図である。

【図13】 演奏データに基づく歌唱合成スコアの形成状況及び歌唱音声の合成状況を示す図である。

30

【図14】 図13の音韻トラックにおける種々のアイテムを説明するための図である。

【図15】 図13の遷移トラックにおける種々のアイテムを説明するための図である。

【図16】 図13のビブラートトラックにおける種々のアイテムを説明するための図である。

【図17】 演奏データ受信・歌唱合成スコア形成処理を示すフローチャートである。

【図18】 歌唱合成スコア形成処理を示すフローチャートである。

【図19】 管理データ作成処理を示すフローチャートである。

【図20】 Event State = Transitionの場合の管理データ作成処理を説明するための図である。

40

【図21】 Event State = Attackの場合の管理データ作成処理を説明するための図である。

【図22】 音韻トラック形成処理を示すフローチャートである。

【図23】 音韻遷移時間長獲得処理を示すフローチャートである。

【図24】 Silence歌唱長算出処理を示すフローチャートである。

【図25】 図24の処理においてConsonant伸縮率が1より大きい場合のConsonant歌唱長算出処理を説明するための図である。

【図26】 図24の処理においてConsonant伸縮率が1より小さい場合のConsonant歌唱長算出処理を説明するための図である。

【図27】 Silence歌唱長の算出例を示す図である。

50

【図28】先行Vowel歌唱長算出処理を示すフローチャートである。

【図29】図28の処理においてConsonant伸縮率が1より大きい場合のConsonant歌唱長算出処理を説明するための図である。

【図30】図28の処理においてConsonant伸縮率が1より小さい場合のConsonant歌唱長算出処理を説明するための図である。

【図31】先行Vowel歌唱長算出例を示す図である。

【図32】Vowel歌唱長算出処理を示すフローチャートである。

【図33】Vowel歌唱長算出例を示す図である。

【図34】遷移トラック形成処理を示すフローチャートである。

【図35】NONEn, NONEs遷移時間長の算出例を示す図である。 10

【図36】pNONEs, NONEs遷移時間長の算出例を示す図である。

【図37】ビブラートトラック形成処理を示すフローチャートである。

【図38】ビブラートトラックの形成例を示す図である。

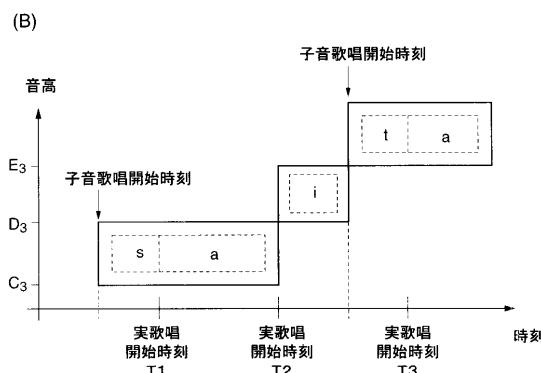
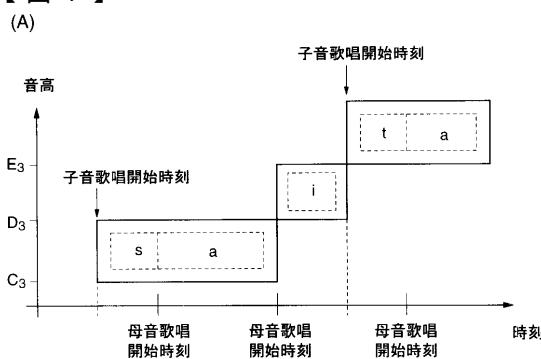
【図39】Silence歌唱長算出の変形例を示す図である。

【図40】人歌唱と、従来の歌唱合成とで歌唱開始時刻を対比して示す図である。

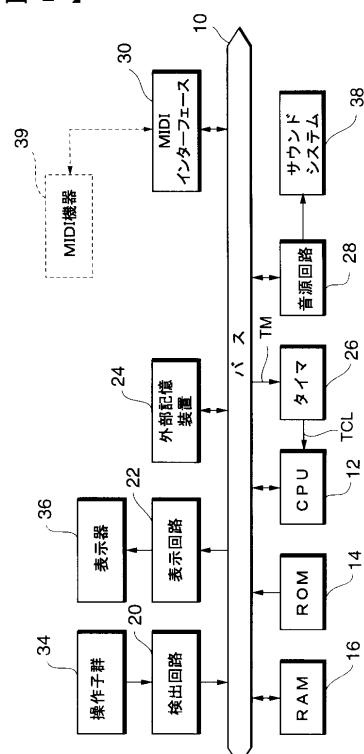
【符号の説明】

10 : バス、12 : CPU、14 : ROM、16 : RAM、20 : 検出回路、22 : 表示回路、24 : 外部記憶装置、26 : タイマ、28 : 音源回路、30 : MIDIインターフェース、34 : 操作子群、36 : 表示器、38 : サウンドシステム、39 : MIDI機器。
20

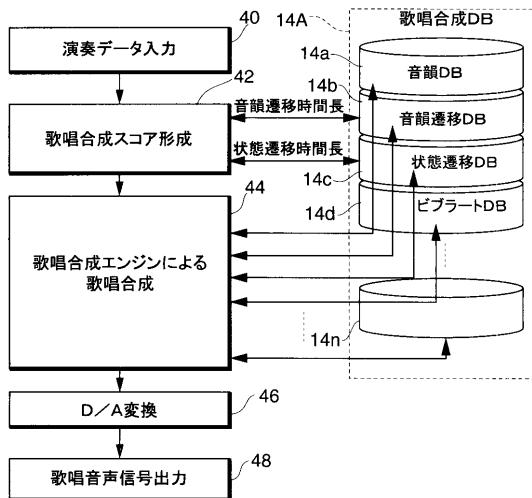
【図1】



【図2】



【 四 3 】



【 図 4 】

演奏データ内の情報		
演奏情報	情報	解説
ノート情報	Note On	実歌唱開始時刻
	Duration	実歌唱長
	Pitch	歌唱音高
音韻トラック情報	PhU	歌唱音韻
	Consonant Modification	歌唱子音伸縮率
遷移トラック情報	Attack Type	歌唱アタックタイプ
	Attack Rate	歌唱アタック伸縮率
	Release Type	歌唱リリースタイプ
	Release Rate	歌唱リリース伸縮率
	Note Transition Type	歌唱ノート遷移タイプ
	Note Transition Rate	歌唱ノート遷移伸縮率
ビブラートトラック情報	Vibrato Number	演奏データ中のビブラートイベント数
	Vibrato Delay1	1番目ビブラート遅延時間
	Vibrato Duration1	1番目ビブラート継続時間
	Vibrato Type1	1番目ビブラートタイプ
	⋮	⋮
	Vibrato DelayK	K番目ビブラート遅延時間
	Vibrato DurationK	K番目ビブラート継続時間
	Vibrato TypeK	K番目ビブラートタイプ

〔 図 5 〕

【 図 6 】

(A)

音韻遷移DB14b内の情報

音韻遷移時間長

- (a) V_Sil 遷移時間長
- (b) Sil_C 遷移時間長
- (c) C_V 遷移時間長
- (d) Sil_V 遷移時間長
- (e) pV_C 遷移時間長
- (f) pV_V 遷移時間長

(B)

先行音韻	後続音韻	音高	音韻遷移時間長	音源制御情報
a	i	P1	音韻遷移時間長	音源制御情報
		P2	音韻遷移時間長	音源制御情報
		⋮		
	M	P1	⋮	⋮
		⋮		
	⋮	⋮		
	Aspiration	⋮		
i	Sil	⋮		
	a	⋮		
	M	⋮		
	⋮	⋮		
	Aspiration	⋮		
M	Sil	⋮		
	a	⋮		
	i	⋮		
	⋮	⋮		
	Aspiration	⋮		
⋮	Sil	⋮		
	⋮	⋮		
	⋮	⋮		
	⋮	⋮		
Sil	a	⋮		
	i	⋮		
	M	⋮		
	⋮	⋮		
	Aspiration	⋮		

【 図 7 】

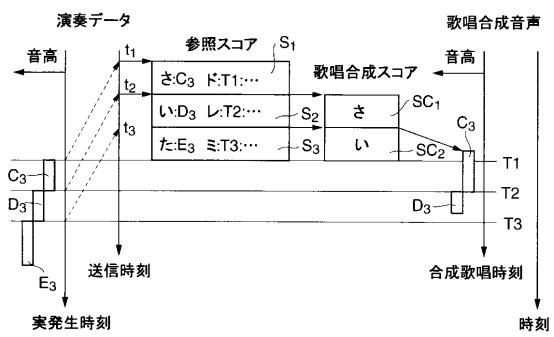
遷移状態		状態タイプ	音韻	音高	状態遷移時間長	音源制御情報
			a	P1 P2 ⋮ P1	状態遷移時間長	音源制御情報
Normal	i		a			
	M		i			
	M		M			
Sexy	M		a			
	M		i			
	M		M			
Attack	M		a			
	M		i			
	M		M			
Sharp	M		a			
	M		i			
	M		M			
Soft	M		a			
	M		i			
	M		M			
Normal	M		a			
	M		i			
	M		M			
NtN	M		a			
	M		i			
	M		M			
Release	M		a			
	M		i			
	M		M			

【図8】

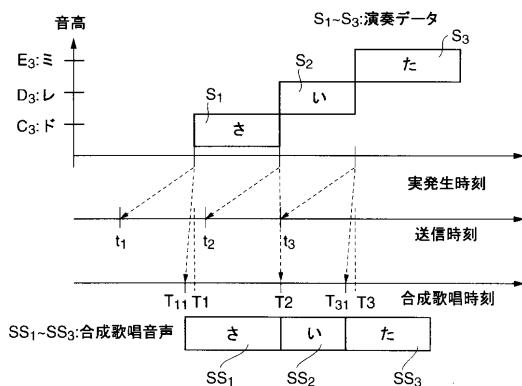
ピラートDB14d内の情報

ピラートタイプ	音韻	音高	音源制御情報
Normal	a	P1 P2 ⋮	音源制御情報
	i	P1 ⋮	⋮
	M	⋮	⋮
	⋮	⋮	⋮
Sexy	a	⋮	⋮
	i	⋮	⋮
	M	⋮	⋮
	⋮	⋮	⋮
Enka	a	⋮	⋮
	i	⋮	⋮
	M	⋮	⋮
	⋮	⋮	⋮

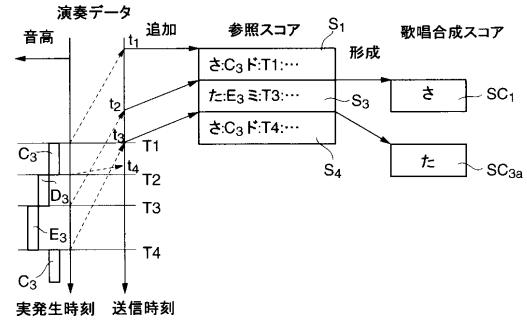
【図10】



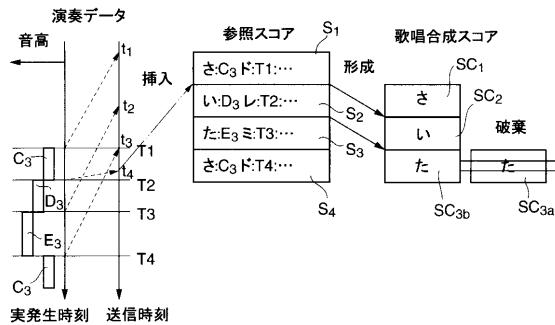
【図9】



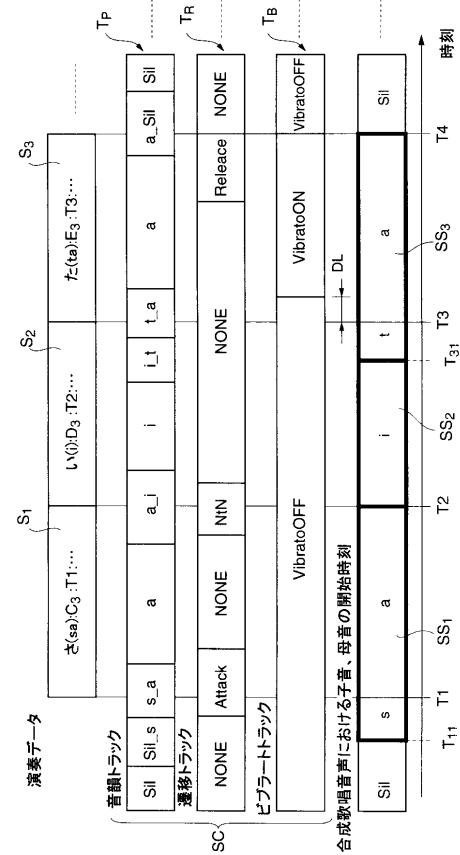
【図11】



【図12】



【図13】



【図14】

音韻トラックTpのアイテム	情報	内容	解説
Sil	Begin Time	T11	開始時刻
	Duration	D11	継続時間
	PhU	Sil	音韻
Sil,s	Begin Time	T12	開始時刻
	Duration	D12	継続時間
	PhU1,PhU2	Sil,s	音韻
s,a	Begin Time	T13	開始時刻
	Duration	D13	継続時間
	PhU1,PhU2	s, a	音韻
a	Begin Time	T14	開始時刻
	Duration	D14	継続時間
	PhU	a	音韻

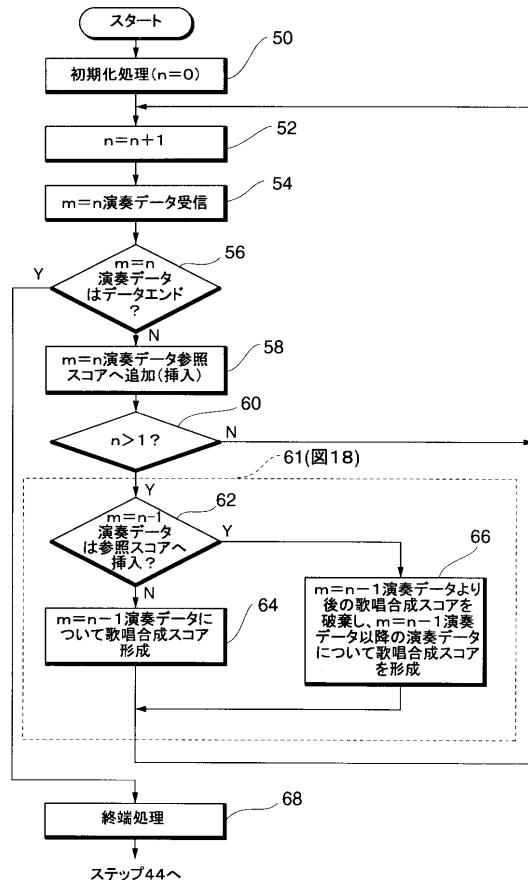
【図16】

ビブラートトラックTbのアイテム	情報	内容	解説
Vibrato OFF	Begin Time	T31	開始時刻
	Duration	D31	継続時間
	Index	OFF	インデックス
Vibrato ON	Begin Time	T32	開始時刻
	Duration	D32	継続時間
	Index	ON	インデックス
Vibrato OFF	Type	Type32	ビブラートのタイプ
	Begin Time	T33	開始時刻
	Duration	D33	継続時間
	Index	OFF	インデックス

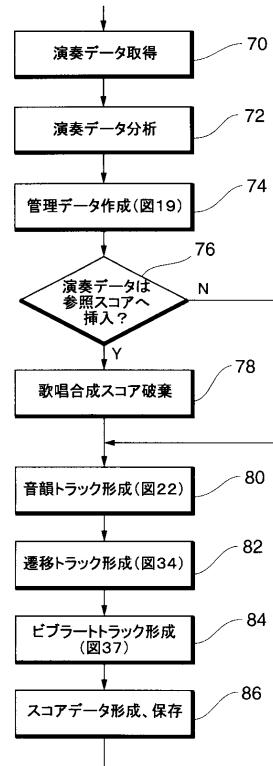
【図15】

遷移トラックTrのアイテム	情報	内容	解説
NONE	Begin Time	T21	開始時刻
	Duration	D21	継続時間
	Index	NONE	遷移インデックス
Attack	Begin Time	T22	開始時刻
	Duration	D22	継続時間
	Index	Attack	遷移インデックス
	Type	Type22	遷移インデックスのタイプ
NONE	Begin Time	T23	開始時刻
	Duration	D23	継続時間
	Index	NONE	遷移インデックス
NtN	Begin Time	T24	開始時刻
	Duration	D24	継続時間
	Index	NtN	遷移インデックス
	Type	Type24	遷移インデックスのタイプ
NONE	Begin Time	T25	開始時刻
	Duration	D25	継続時間
	Index	NONE	遷移インデックス
Release	Begin Time	T26	開始時刻
	Duration	D26	継続時間
	Index	Release	遷移インデックス
	Type	Type26	遷移インデックスのタイプ

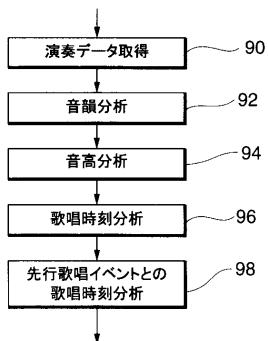
【図17】



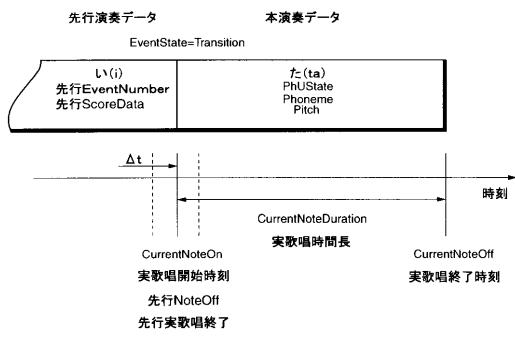
【図18】



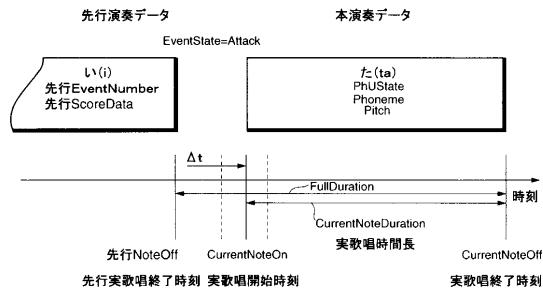
【図19】



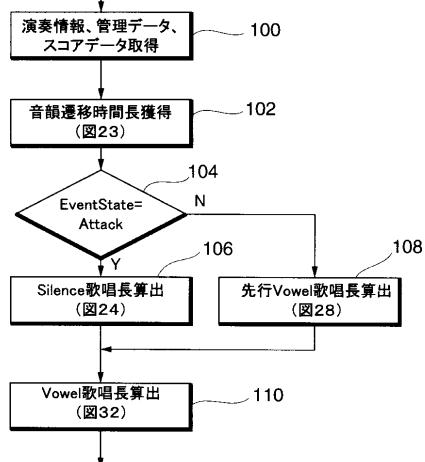
【図20】



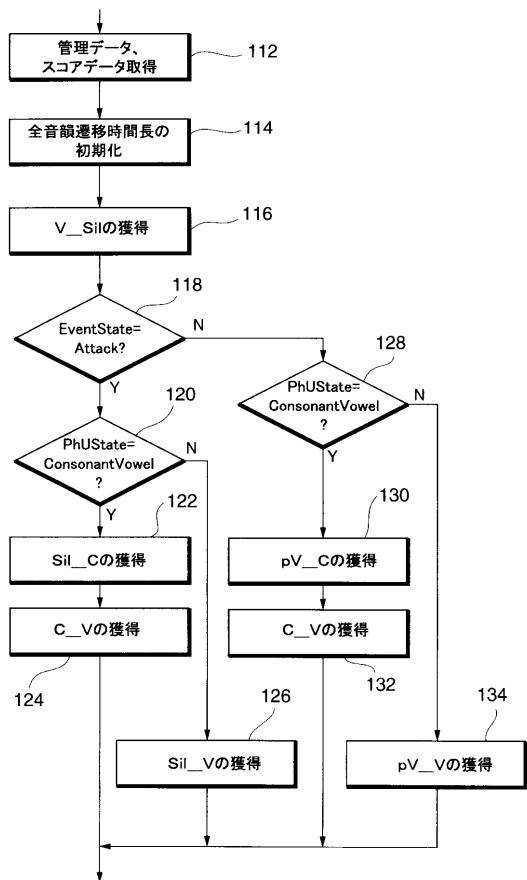
【図21】



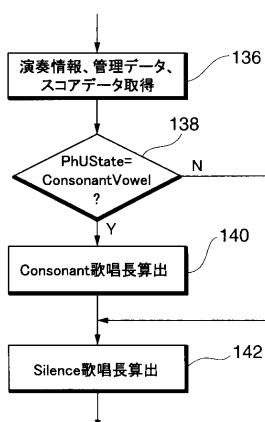
【図22】



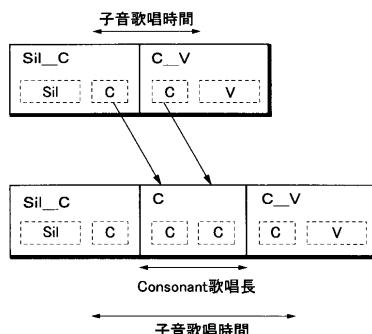
【図23】



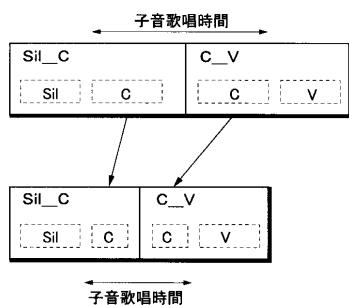
【図24】



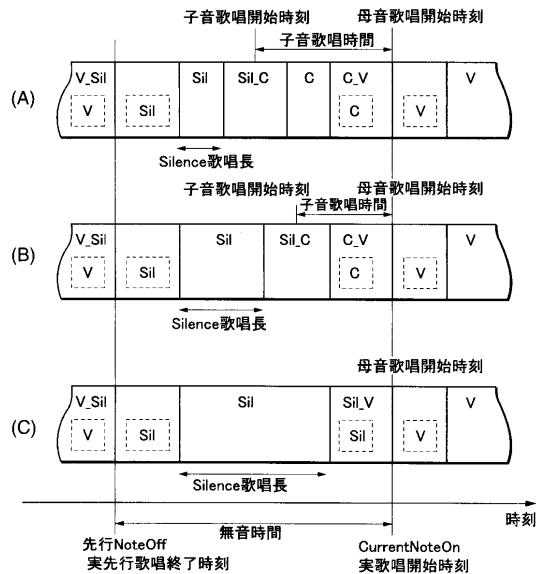
【図25】



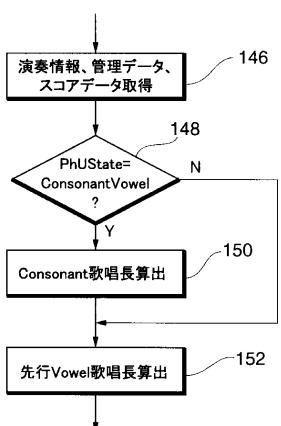
【図26】



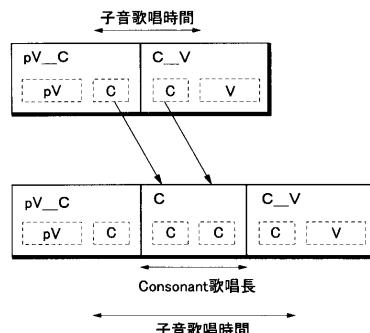
【図27】



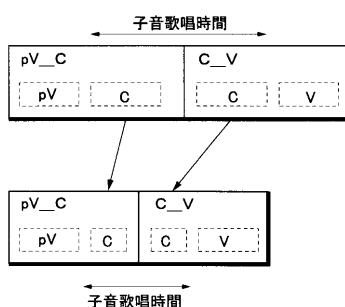
【図28】



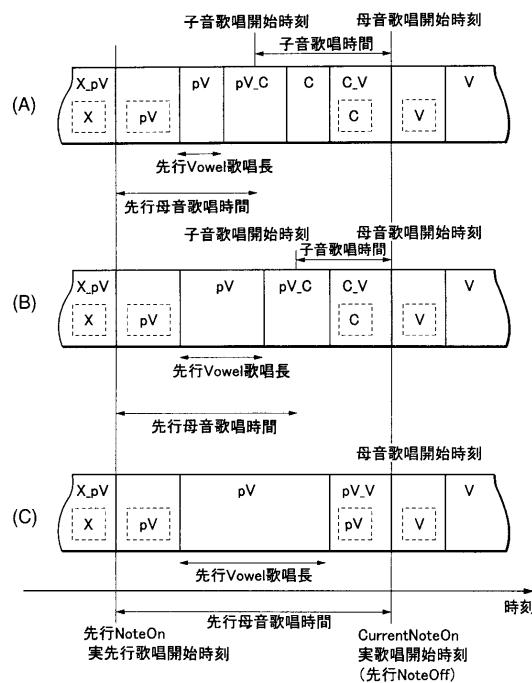
【図29】



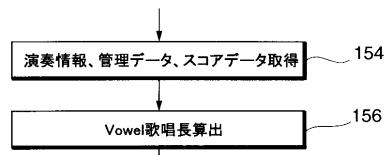
【図30】



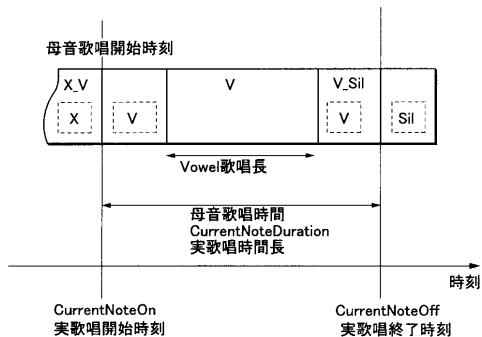
【図31】



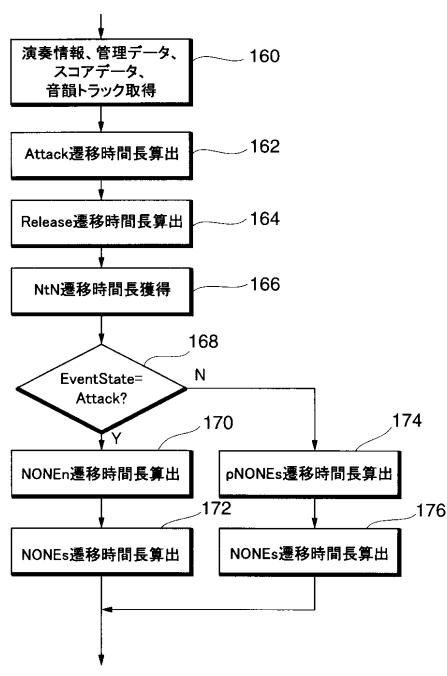
【図32】



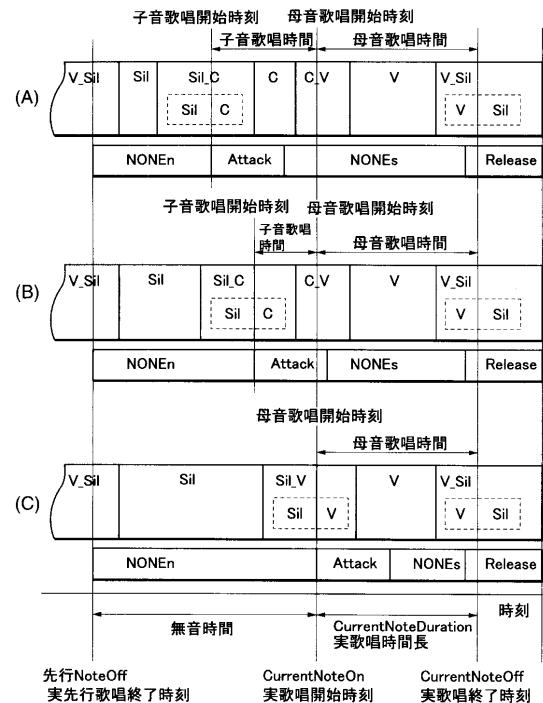
【図33】



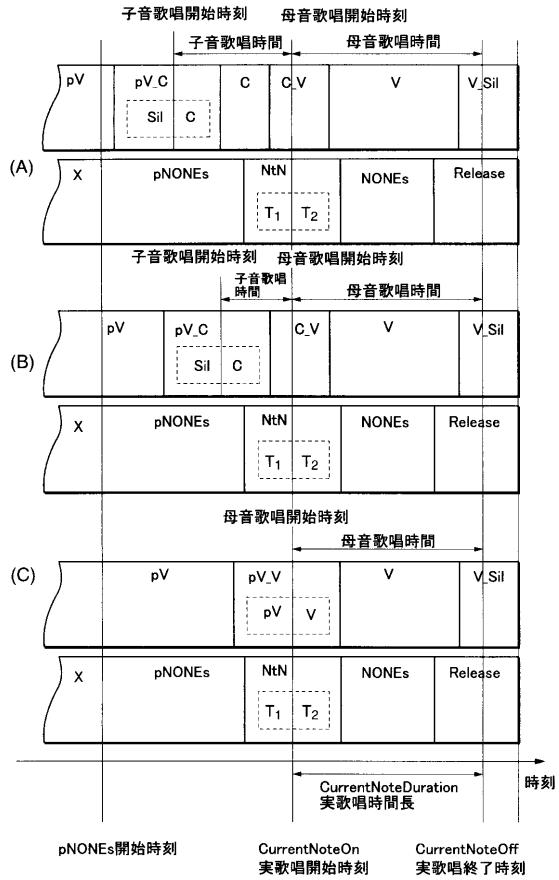
【図34】



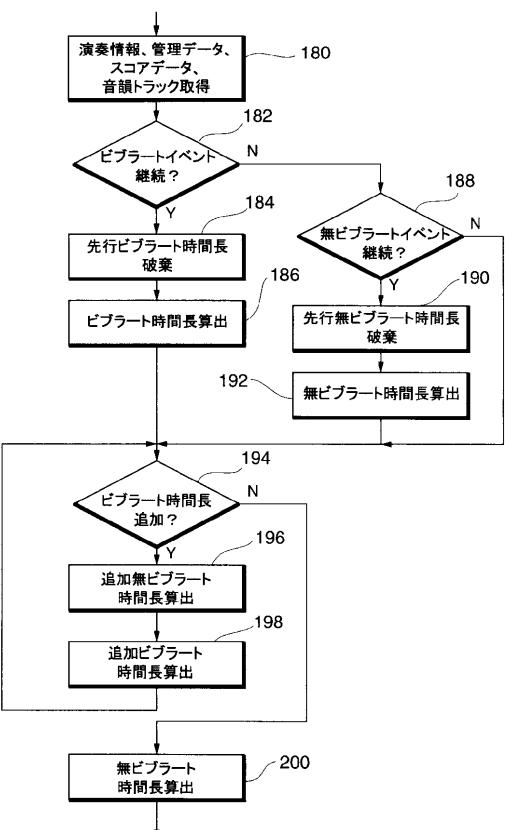
【図35】



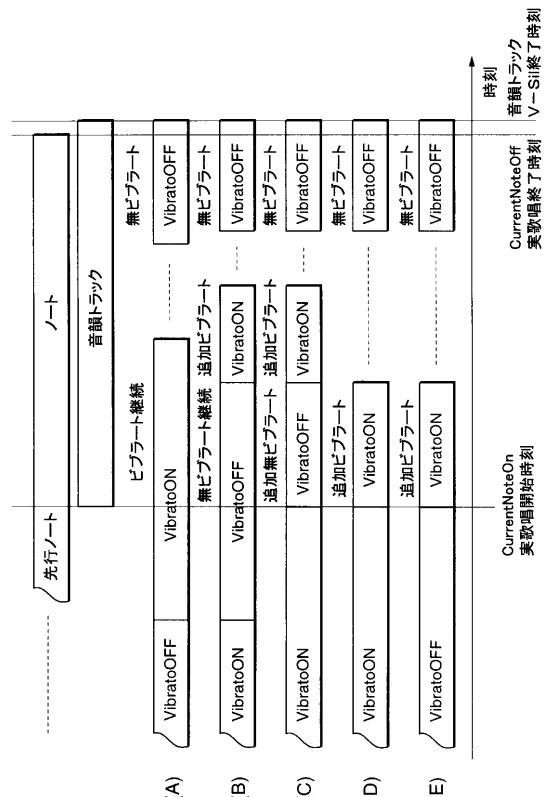
【図36】



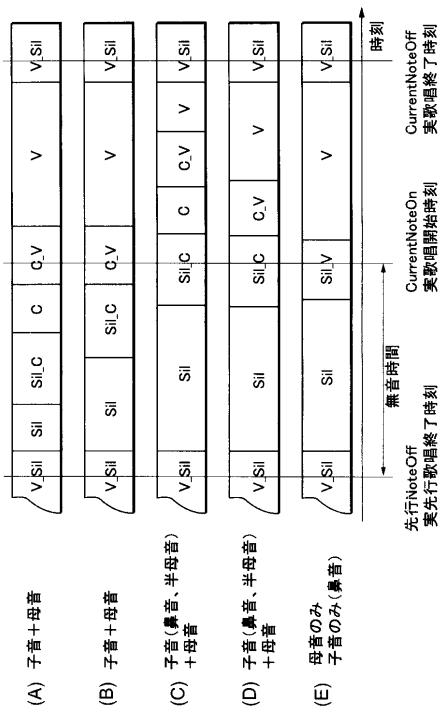
【図37】



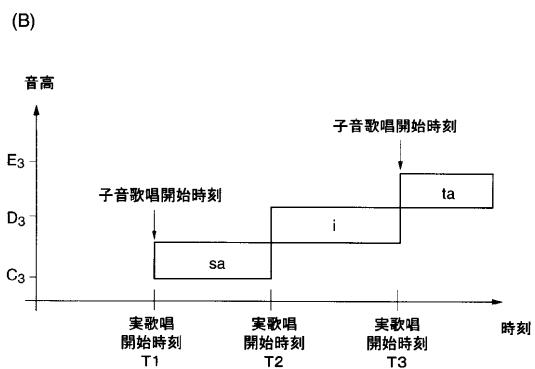
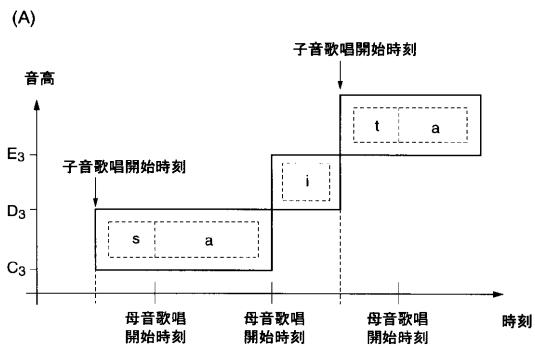
【図38】



【図39】



【図 4 0】



フロントページの続き

審査官 荘原 雄一

(56)参考文献 特開平10-319993(JP,A)
特開平08-248993(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 13/02

G10H 1/00