(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2012-113271
(P2012-113271A)
(43)公開日　平成24年6月14日(2012.6.14)

(51)Int.Cl.			F I			テーマコード（参考）
G10H	1/00	(2006.01)	G10H	1/00	Z	5D082
G10G	1/00	(2006.01)	G10H	1/00	102Z	5D378
			G10G	1/00		

審査請求 未請求　請求項の数 5　OL　（全 14 頁）

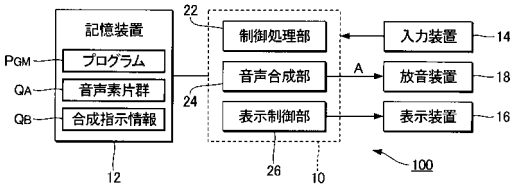
(21)出願番号　特願2010-264643 (P2010-264643)
(22)出願日　平成22年11月29日 (2010.11.29)

(71)出願人　000004075
ヤマハ株式会社
静岡県浜松市中区中沢町１０番１号
(74)代理人　100125689
弁理士　大林　章
(74)代理人　100125335
弁理士　矢代　仁
(74)代理人　100121108
弁理士　高橋　太朗
(72)発明者　久湊　裕司
静岡県浜松市中区中沢町１０番１号　ヤマハ株式会社内

Fターム(参考)　5D082 AA03　AA06　AA24
5D378 MM34　MM35　MM51　MM65　MM94
TT09　TT23　TT24

(54)【発明の名称】音楽情報表示装置

(57)【要約】　（修正有）
【課題】拍点基準で規定された音楽情報の時系列と実時間基準で規定された音楽情報の時系列との時間的な対応を利用者に容易に確認させる。
【解決手段】記憶装置１２は、合成対象音の音高を拍点単位で時系列に指定する合成音情報と、実時間軸上の音響信号の波形を示す波形情報とを記憶する。制御処理部２２は、利用者からの指示に応じてテンポを可変に設定する。表示制御部２６は、合成音情報が指定する合成対象音の時系列を示す対象音画像と音響信号の波形を示す波形画像とを拍点単位の共通の拍点軸のもとで表示させ、制御処理部２２が設定するテンポに応じて波形画像を拍点軸の方向に伸縮する。
【選択図】図１



【特許請求の範囲】
【請求項１】
第１音楽情報を拍点単位で時系列に指定する第１時系列情報を記憶する第１記憶手段と、
第２音楽情報を実時間単位で時系列に指定する第２時系列情報を記憶する第２記憶手段と、
テンポを可変に設定するテンポ設定手段と、
前記第１時系列情報を示す第１時系列画像と前記第２時系列情報を示す第２時系列画像とを拍点単位の共通の拍点軸のもとで表示装置に表示させ、前記テンポ設定手段が設定するテンポに応じて前記第２時系列画像を前記拍点軸の方向に伸縮する表示制御手段と
を具備する音楽情報表示装置。
【請求項２】
前記第２時系列情報は、振幅値を前記第２音楽情報として音響信号の時間波形を指定し、
前記表示制御手段は、前記第２時系列情報から特定される推定拍点に対応する拍点指示子を前記第２時系列画像とともに前記拍点軸のもとで前記表示装置に表示させ、前記テンポ設定手段が設定するテンポに応じて前記各拍点指示子を前記拍点軸の方向に移動させる
請求項１の音楽情報表示装置。
【請求項３】
前記第２時系列情報から特定される推定拍点が前記拍点軸上の各拍点に合致するように、前記表示制御手段が前記第２時系列画像を前記拍点軸の方向に伸縮するとともに前記テンポ設定手段がテンポを設定する
請求項２の音楽情報表示装置。
【請求項４】
前記第２音楽情報の再生速度の変更の許可／禁止を選択し、前記再生速度の変更を許可した場合に、前記テンポ設定手段が設定するテンポに応じて前記第２音楽情報の時系列が時間軸上で伸縮するように前記第２時系列情報を変更し、前記再生速度の変更を禁止した場合には前記第２時系列情報を変更しない制御手段
を具備する請求項１の音楽情報表示装置。
【請求項５】
前記制御手段は、前記第２音楽情報の再生速度の変更の許可／禁止を利用者からの指示に応じて選択する
請求項４の音楽情報表示装置。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、音楽に関連する情報（音楽情報）の時系列を表示する技術に関する。
【背景技術】
【０００２】
音声合成の対象となる音声（以下「合成対象音」という）の時系列を生成および編集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、拍点を単位とする横方向の時間軸（以下では「拍点軸」という）と音高を示す縦方向の音高軸とが規定されたピアノロール型の編集画面を参照して利用者が各合成対象音を編集する技術が開示されている。合成対象音を示す音指示子（ノートバー）を編集画面に適宜に配置することで、利用者は、拍点を単位（基準）として所望の合成対象音を指定することが可能である。
【先行技術文献】
【特許文献】
【０００３】
【特許文献１】特開２００６－２５９７６８号公報

【発明の概要】
【発明が解決しようとする課題】
【０００４】
ところで、編集画面で指定された各合成対象音の時系列は、実時間（例えば秒）を単位とする時間軸（以下「実時間軸」という）のもとで生成された他の音楽情報の時系列と並列に処理され得る。例えば、実時間軸での時間波形として収録された伴奏音を合成対象音の時系列と並列に再生することが可能である。また、合成対象音毎に音指示子が配置された編集画面と伴奏音の時間波形とが並列に表示されれば、例えば伴奏音のうち音量が瞬間的に増加する時点（すなわち拍点）に合致するように合成対象音を配置するといった調整が実現されて便利である。
【０００５】
しかし、各合成対象音の時系列が拍点軸のもとで表示され、伴奏音の時間波形が実時間軸のもとで表示される構成では、両者間の時間的な対応（同期）を利用者が把握することが困難であるという問題がある。なお、以上の説明では伴奏音の時間波形を実時間上の時系列情報として例示したが、伴奏音の時間波形以外の実時間上の時系列情報を拍点軸上の時系列情報と並列に処理する場合にも、同様の問題が発生し得る。以上の事情を背景として、本発明は、拍点基準で規定された音楽情報の時系列と実時間基準で規定された音楽情報の時系列との時間的な対応を利用者に容易に確認させることを目的とする。
【課題を解決するための手段】
【０００６】
以上の課題を解決するために本発明が採用する手段を説明する。なお、本発明の理解を容易にするために、以下の説明では、本発明の要素と後述の実施形態の要素との対応を括弧書で付記するが、本発明の範囲を実施形態の例示に限定する趣旨ではない。
【０００７】
本発明の音楽情報表示装置は、第１音楽情報を拍点単位で時系列に指定する第１時系列情報（例えば合成音情報ＤＳ）を記憶する第１記憶手段（例えば記憶装置１２）と、第２音楽情報を実時間単位で時系列に指定する第２時系列情報（例えば波形情報ＤＷや合成音情報ＤＲ）を記憶する第２記憶手段（例えば記憶装置１２）と、テンポを可変に設定するテンポ設定手段（例えば制御処理部２２）と、第１時系列情報を示す第１時系列画像（例えば対象音画像４２）と第２時系列情報を示す第２時系列画像（例えば波形画像４４や音素列画像６２Bや特徴量画像６４B）とを拍点単位の共通の拍点軸のもとで表示装置に表示させ、テンポ設定手段が設定するテンポに応じて第２時系列画像を拍点軸の方向に伸縮する表示制御手段（例えば表示制御部２６）とを具備する。以上の構成によれば、第１時系列画像と第２時系列画像とが共通の拍点軸のもとで表示されるとともに第２時系列画像がテンポに応じて拍点軸の方向に伸縮されるから、拍点基準で規定される第１音楽情報の時系列と実時間基準で規定される第２音楽情報の時系列との時間的な対応を利用者が容易に把握できるという利点がある。
【０００８】
本発明の好適な態様において、第２時系列情報（例えば波形情報ＤＷ）は、振幅値を第２音楽情報として音響信号（例えば音響信号ＡＷ）の時間波形を指定し、表示制御手段は、第２時系列情報から特定される推定拍点に対応する拍点指示子（例えば拍点指示子４４２）を第２時系列画像とともに拍点軸のもとで表示装置に表示させ、テンポ設定手段が設定するテンポに応じて各拍点指示子を拍点軸の方向に移動させる。以上の態様では、第２時系列情報から特定される推定拍点に対応する拍点指示子が拍点軸のもとで表示されるとともに各拍点指示子がテンポに応じて拍点軸の方向に移動されるから、第１時系列情報が示す第１音楽情報の時系列と第２時系列情報が指定する音響信号の推定拍点との時間的な対応を利用者が容易に把握できるという利点がある。
【０００９】
本発明の好適な態様において、第２時系列情報から特定される推定拍点が拍点軸上の各拍点に合致するように、表示制御手段が第２時系列画像を拍点軸の方向に伸縮するととも

にテンポ設定手段がテンポを設定する。以上の態様では、推定拍点が拍点軸上の各拍点に合致するようにテンポが設定されるから、例えば第１時系列画像と第２時系列画像との対応を確認しながら第１音楽情報の時系列と第２音楽情報の時系列とが同期するように利用者が手動でテンポを調整する構成と比較して、利用者の負担が軽減されるという利点がある。なお、以上の態様の具体例は例えば第４実施形態として後述される。

【００１０】

本発明の好適な態様に係る音楽情報表示装置は、第２音楽情報の再生速度（例えば再生速度Ｖ2）の変更の許可／禁止を選択し、再生速度の変更を許可した場合に、テンポ設定手段が設定するテンポに応じて第２音楽情報の時系列が時間軸上で伸縮するように第２時系列情報を変更し、再生速度の変更を禁止した場合には第２時系列情報を変更しない制御手段（例えば制御処理部２２）を具備する。以上の態様においては、第２音楽情報の再生速度の変更の許否が選択されるから、例えば第２音楽情報の時系列を拍点に同期させる必要がない場合にはテンポに応じた再生速度の変更を禁止し、第２音楽情報の時系列を拍点に同期させる必要がある場合にはテンポに応じた再生速度の変更を許可するといった制御が可能である。なお、以上の態様の具体例は例えば第２実施形態として後述される。

【００１１】

本発明の好適な態様において、制御手段は、第２音楽情報の再生速度の変更の許可／禁止を利用者からの指示に応じて選択する。以上の態様では、第２音楽情報の再生速度の変更の許可／禁止が利用者からの指示に応じて選択されるから、第２音楽情報の時系列を拍点に同期させるか否かを利用者の意図に応じて変更できるという利点がある。

【００１２】

以上の各態様に係る音楽情報表示装置は、音楽情報の表示制御に専用されるＤＳＰ（Digital Signal Processor）などのハードウェア（電子回路）によって実現されるほか、ＣＰＵ（Central Processing Unit）等の汎用の演算処理装置とプログラムとの協働によっても実現される。本発明のプログラムは、第１音楽情報を拍点単位で時系列に指定する第１時系列情報を記憶する第１記憶手段と、第２音楽情報を実時間単位で時系列に指定する第２時系列情報を記憶する第２記憶手段とを具備するコンピュータに、テンポを可変に設定するテンポ設定処理と、第１時系列情報を示す第１時系列画像と第２時系列情報を示す第２時系列画像とを拍点単位の共通の拍点軸のもとで表示装置に表示させ、テンポ設定処理で設定したテンポに応じて第２時系列画像を拍点軸の方向に伸縮する表示制御処理とを実行させる。以上のプログラムによれば、本発明に係る音楽情報表示装置と同様の作用および効果が奏される。本発明のプログラムは、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利用者に提供されてコンピュータにインストールされるほか、通信網を介した配信の形態でサーバ装置から提供されてコンピュータにインストールされる。

【図面の簡単な説明】

【００１３】

【図１】本発明の第１実施形態に係る音声合成装置のブロック図である。

【図２】合成指示情報の模式図である。

【図３】編集画面の模式図である。

【図４】操作画面の模式図である。

【図５】テンポに応じた波形画像の伸縮の説明図である。

【図６】対象音編集画面の模式図である。

【図７】第２実施形態における合成指示情報の模式図である。

【図８】第２実施形態における編集画面の模式図である。

【図９】第２実施形態における操作画面の模式図である。

【図１０】対象音編集画面（実時間基準）での音素列画像および特徴量画像の伸縮の説明図である。

【図１１】編集画面（拍点基準）での音素列画像および特徴量画像の伸縮の説明図である。

【図１２】第３実施形態における操作画面の模式図である。

【図１３】第３実施形態におけるテンポに応じた対象音画像の伸縮の説明図である。
【図１４】第４実施形態における操作画面の模式図である。
【発明を実施するための形態】
【００１４】
＜Ａ：第１実施形態＞
図１は、本発明の第１実施形態の音声合成装置１００のブロック図である。音声合成装置１００は、音響信号Ａを合成する音響処理装置であり、図１に示すように、演算処理装置１０と記憶装置１２と入力装置１４と表示装置１６と放音装置１８とを具備するコンピュータシステムで実現される。入力装置１４（例えばマウスやキーボード）は、利用者からの指示を受付ける。表示装置１６（例えば液晶表示装置）は、演算処理装置１０から指示された画像を表示する。放音装置１８（例えばスピーカやヘッドホン）は、演算処理装置１０が生成した音響信号Ａに応じた音響を再生する。
【００１５】
記憶装置１２は、演算処理装置１０が実行するプログラムＰＧＭや演算処理装置１０が使用する情報（例えば音声素片群ＱＡ，合成指示情報ＱＢ）を記憶する。半導体記録媒体や磁気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が記憶装置１２として任意に採用され得る。
【００１６】
音声素片群ＱＡは、相異なる音声素片に対応する複数の素片情報（例えば音声素片の波形のサンプル系列）で構成されて音響信号Ａの素材として利用される音声合成用ライブラリである。音声素片は、言語上の意味の区別の最小単位（例えば母音や子音）である１個の音素、または複数の音素を連結した音素連鎖である。合成指示情報ＱＢは、音響信号Ａが示す音響の内容を指示する情報であり、図２に示すように、合成音情報ＤＳと波形情報ＤＷとテンポ情報ＤＴとを含んで構成される。
【００１７】
合成音情報ＤＳは、合成対象音の時系列を指定する時系列情報（スコアデータ）である。図２に示すように、合成音情報ＤＳは、相異なる合成対象音に対応する複数の単位情報ＵＳの時系列で構成される。各単位情報ＵＳは、合成対象音の音高（ノートナンバ）ｓ1と発音文字ｓ2と発音期間ｓ3とを指定する。発音文字ｓ2は、合成対象音の発音の音節（文字）である。
【００１８】
発音期間ｓ3は、合成対象音の発音の始点と終点とで規定される。発音の始点および終点の各々は、秒単位等の実時間ではなく拍点を基準として（あるいは複数の拍点を含む小節を基準として）指定される。具体的には、発音期間ｓ3の各端点（始点,終点）は、小節番号σ1と拍点番号σ2と区間番号σ3との組合せで指定される。区間番号σ3は、相前後する２個の拍点の間隔（以下「拍点区間」という）を区分した複数の単位区間の何れかを識別する番号である。すなわち、第σ1番目の小節のうち第σ2番目の拍点から合計σ3個の単位区間が経過した時点が発音期間ｓ3の端点として指定される。以上の説明から理解されるように、合成音情報ＤＳは、合成対象音の音高ｓ1や発音文字ｓ2を拍点基準で（すなわち実時間ではなく拍点や小節を基準として）時系列に指定する。
【００１９】
図２の波形情報ＤＷは、楽曲の伴奏音等の音楽を収録した音響信号ＡＷの時間波形を指定する。例えば利用者が選択した音響信号ＡＷが光記録媒体や半導体記録媒体等の各種の記録媒体から波形情報ＤＷとして記憶装置１２に転送される。波形情報ＤＷは、事前に収録された音響信号ＡＷの振幅値を実時間軸上の所定の周期毎に抽出した多数のサンプルの時系列で構成される。すなわち、波形情報ＤＷは、音響信号ＡＷの振幅値を実時間基準で（すなわち実時間を単位（例えば秒単位）として）時系列に指定する。
【００２０】
図２のテンポ情報ＤＴは、拍点が到来するテンポ（速度）を時系列に指定する情報であり、テンポの指定単位となる速度調整区間毎に単位情報（テンポチェンジメッセージ）Ｕ

Tを含んで構成される。各単位情報UTは、テンポを変更する時点ｔ1（すなわち速度調整区間の始点）と時点ｔ1以後のテンポｔ2とを指定する。速度調整区間の開始の時点ｔ1は、前述の発音期間ｓ3の端点と同様に拍点基準で指定される。すなわち、例えば小節番号σ1と拍点番号σ2と区間番号σ3との組合せで各単位情報UTの時点ｔ1が指定される。以上が合成指示情報QBの具体的な内容である。

【００２１】

図１の演算処理装置１０は、記憶装置１２に格納されたプログラムPGMの実行で、音響信号Aの生成に必要な複数の機能（制御処理部２２，音声合成部２４，表示制御部２６）を実現する。なお、演算処理装置１０の各機能を専用の電子回路（ＤＳＰ）で実現した構成や、演算処理装置１０の各機能を複数の集積回路に分散した構成も採用され得る。制御処理部２２は、入力装置１４に対する利用者からの指示に応じて合成指示情報QB（合成音情報DS，波形情報DW，テンポ情報DT）を生成および更新する（詳細は後述）。

【００２２】

音声合成部２４は、記憶装置１２に格納された音声素片群QAおよび合成指示情報QBを利用して音響信号Aを生成する。第１実施形態の音声合成部２４は、図２に示すように、合成音情報DSに応じた音響信号ASと波形情報DWが示す音響信号AWとの混合で音響信号Aを生成する。具体的には、音声合成部２４は、記憶装置１２に格納された音声素片群QAのうち合成音情報DSの各単位情報USが指定する発音文字ｓ2に対応した素片情報を、テンポ情報DTの単位情報UTが指定するテンポｔ2と当該単位情報USの発音期間ｓ3とに応じた時間長に伸縮するとともに当該単位情報USが指定する音高ｓ1に調整し、調整後の各素片情報を相互に連結することで音響信号ASを生成する。そして、音声合成部２４は、音響信号ASと波形情報DWが指定する音響信号AWとを混合することで音響信号Aを生成する。音響信号Aは、放音装置１８に供給されて音響として再生される。

【００２３】

図１の表示制御部２６は、合成対象音の編集時に利用者が視認する図３の編集画面３０を表示装置１６に表示させる。図３に示すように、拍点を示す拍点軸Xbが編集画面３０に規定される。拍点軸Xb上には複数の拍点が所定の間隔（等間隔）で離散的に設定され、拍点を示す直線状の基準指示子４０が拍点軸Xb上の各拍点の位置に配置される。図３に示すように、編集画面３０は、第１領域３１と第２領域３２と第３領域３３とを含んで構成される。

【００２４】

第１領域３１には、拍点軸Xbに直交する縦方向の音高軸Xpが規定される。表示制御部２６は、合成対象音の時系列を表現する対象音画像４２を第１領域３１内に配置する。対象音画像４２は、各合成対象音を図形化した複数の音指示子４２２を拍点軸Xbの方向に時系列に配列したピアノロール型の画像である。音高軸Xpの方向における音指示子４２２の位置は合成対象音の音高に応じて可変に設定される。また、拍点軸Xbの方向における音指示子４２２の位置（例えば音指示子４２２の右端部の位置）は合成対象音の発音の時点（例えば発音期間ｓ3の始点）に応じて可変に設定され、拍点軸Xbの方向における音指示子４２２の長さは合成対象音の発音期間ｓ3の長短に応じて可変に設定される。

【００２５】

利用者は、編集画面３０を確認しながら入力装置１４を適宜に操作することで、合成対象音の編集を指示することが可能である。例えば、第１領域３１内の任意の位置に対する音指示子４２２の追加や既存の音指示子４２２の削除、または音指示子４２２の発音文字ｓ2の指定または変更等の編集が利用者から指示される。表示制御部２６は、第１領域３１に対する利用者からの指示に応じて対象音画像４２を随時に更新する。

【００２６】

制御処理部２２は、第１領域３１に配置された音指示子４２２毎に合成音情報DSの各単位情報USを生成して記憶装置１２に格納する。すなわち、制御処理部２２は、音高軸Xpの方向における音指示子４２２の位置に応じた音高ｓ1と、音指示子４２２に指定された発音文字ｓ2と、拍点軸Xbの方向における音指示子４２２の位置や長さに応じた発音期

間ｓ3とを含む単位情報USを生成して利用者からの指示に応じて随時に更新する。

【００２７】

表示制御部２６は、波形情報DWが示す音響信号AWの時間波形の画像（以下「波形画像」という）４４を第２領域３２に配置する。波形情報DWは前述のように実時間基準で規定されるが、第２領域３２の波形画像４４は、実時間軸Ｘｔではなくて、第１領域３１内の対象音画像４２と共通の拍点軸Ｘｂのもとで表示される。また、表示制御部２６は、音響信号AWの拍点（以下「推定拍点」という）を波形情報DWから推定し、各推定拍点を表現する拍点指示子４４２を波形画像４４とともに第２領域３２に配置する。図３の各拍点指示子４４２は、各推定拍点の位置に配置された直線状の図形である。なお、音響信号AWの拍点の推定には公知の技術が任意に採用される。

【００２８】

表示制御部２６は、音声合成のテンポｔ2の時系列を第３領域３３に表示する。具体的には、拍点軸Ｘｂ上に区画された速度調整区間４６毎にテンポｔ2が指定される。合成音情報DSの編集が開始された直後の段階では拍点軸Ｘｂの全区間を速度調整区間４６としてテンポｔ2が初期値に設定され、以下に説明するように、利用者が拍点軸Ｘｂ上に任意に区画した速度調整区間４６毎にテンポｔ2が可変に設定される。

【００２９】

利用者は、入力装置１４に対する所定の操作（例えば第３領域３３のダブルクリック）により、拍点軸Ｘｂ上の任意の時点ｔ1を速度調整区間４６の始点として指定したうえでテンポｔ2の変更を指示する。利用者がテンポｔ2の変更を指示すると、表示制御部２６は、利用者がテンポｔ2を指定するための図４の操作画面５０（ダイアログボックス）を表示装置１６に表示させる。

【００３０】

操作画面５０には指定領域５２が配置される。指定領域５２にはテンポｔ2（図４では「125」）が表示される。利用者は、入力装置１４に対する所定の操作で指定領域５２内のテンポｔ2の変更を指示すること（増加／減少の指示または数値の直接入力）が可能である。制御処理部２２は、利用者が選択した速度調整区間４６に対応する単位情報UTのテンポｔ2を、指定領域５２に対する利用者からの指示に応じて可変に設定する。

【００３１】

拍点軸Ｘｂは各拍点が等間隔で配列するように設定されるから、テンポｔ2が変化した場合でも、拍点軸Ｘｂ上の各拍点の間隔（相互に隣合う各基準指示子４０の間隔）は変化しない。また、合成音情報DSは拍点基準で合成対象音の時系列を指定するから、テンポｔ2が変化した場合でも対象音画像４２は変化しない（拍点軸Ｘｂの方向に伸縮しない）。他方、各拍点の間隔である拍点区間δの実時間長（例えば秒単位の時間長）はテンポｔ2に応じて変化する。例えばテンポｔ2が速いほど１個の拍点区間δの実時間長は短くなる。前述のように波形情報DWは実時間基準で音響信号AWを規定するから、表示制御部２６は、制御処理部２２が設定するテンポｔ2に応じて第２領域３２内の波形画像４４を拍点軸Ｘｂの方向に伸縮させ、波形画像４４の伸縮とともに各拍点指示子４４２を拍点軸Ｘｂの方向に移動させる。

【００３２】

例えば、テンポｔ2が速いほど各拍点区間δの実時間長は短くなるから、図５に矢印ａ1で示すように、表示制御部２６は、速度調整区間４６内のテンポｔ2が上昇するほど、その速度調整区間４６に対応する波形画像４４を拍点軸Ｘｂの方向に伸長させる。すなわち、波形画像４４の各推定拍点を示す拍点指示子４４２の間隔が増加する（拍点区間δ内の推定拍点の個数が減少する）。また、速度調整区間４６内のテンポｔ2が低下した場合（拍点区間δの実時間長が長くなる場合）、図５に矢印ａ2で示すように、表示制御部２６は、その速度調整区間４６に対応する波形画像４４を拍点軸Ｘｂの方向に縮小する。

【００３３】

操作画面５０に対する操作に応じたテンポｔ2の変更と拍点軸Ｘｂの方向の波形画像４４の伸縮とは並列に実行される。すなわち、利用者からの指示に応じて制御処理部２２がテ

ンポｔ2を変更（増減／数値指定）するたびに、テンポｔ2の変更に応じた波形画像４４の伸縮が逐次的に実行される。したがって、利用者は、波形画像４４の伸縮（各拍点指示子４４２の移動）を視覚的に確認しながらテンポｔ2を変更することが可能である。具体的には、波形画像４４の各拍点指示子４４２と拍点軸Ｘb上の各拍点に対応する基準指示子４０とが拍点軸Ｘb上で合致するように速度調整区間４６毎のテンポｔ2を設定することで、合成音情報ＤＳに応じて生成される音響信号ＡＳの拍点と波形情報ＤＷが指定する音響信号ＡＷの拍点とを同期させることが可能である。すなわち、拍点基準の合成音情報ＤＳから生成される音響信号ＡＳと実時間基準の波形情報ＤＷが指定する音響信号ＡＷとの時間的な対応を利用者が容易に確認できるという利点がある。

【００３４】

＜Ｂ：第２実施形態＞

本発明の第２実施形態を以下に説明する。第２実施形態では、第１実施形態と同様に拍点基準で規定される合成対象音（以下では特に「第１合成対象音」という）の時系列のほか、実時間基準で規定される合成対象音（以下「第２合成対象音」という）の時系列について音響信号Ａが生成される。なお、以下に例示する各態様において作用や機能が第１実施形態と同等である要素については、以上の説明で参照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。

【００３５】

利用者は、入力装置１４に対する所定の操作で第２合成対象音の編集を指示することが可能である。第２合成対象音の編集が利用者から指示されると、表示制御部２６は、図６の対象音編集画面６０を表示装置１６に表示させる。

【００３６】

図６に示すように、対象音編集画面６０には、音素列画像６２Tと特徴量画像６４Tとが共通の実時間軸Ｘｔ（横軸）のもとで配置される。音素列画像６２Tは、第２合成対象音を構成する音素を実時間軸Ｘｔに沿って時系列に配列した画像である。他方、特徴量画像６４Tは、第２合成対象音の特徴量（以下の例示では音高）の時間変化を表現する画像である。図６では、実時間軸Ｘｔに沿って時系列に配列された複数の編集点α（ブレークポイント）を相互に連結した折線を特徴量画像６４Tとして例示した。

【００３７】

利用者は、対象音編集画面６０を確認しながら入力装置１４を適宜に操作することで第２合成対象音の編集を指示することが可能である。具体的には、音素列画像６２Tを構成する各音素の追加や変更等の編集や、任意の位置に対する編集点αの追加または既存の編集点αの移動および削除等の特徴量画像６４Tの編集が利用者から指示される。表示制御部２６は、利用者からの指示に応じて対象音編集画面６０の音素列画像６２Tおよび特徴量画像６４Tを更新する。

【００３８】

図７は、第２実施形態の合成指示情報ＱＢの模式図である。図７に示すように、第２実施形態の合成指示情報ＱＢは、第１実施形態と同様の要素（合成音情報ＤＳ，波形情報ＤＷ，テンポ情報ＤＴ）に合成音情報ＤＲを追加した構造である。合成音情報ＤＲは、第２合成対象音の時系列を指定する情報であり、音素列画像６２Tに対応する音素列情報ＲＡと特徴量画像６４Tに対応する特徴量情報ＲＢとを含んで構成される。

【００３９】

音素列情報ＲＡは、第２合成対象音の各音素に対応する単位情報ＵＲａの時系列である。単位情報ＵＲａは、音素の識別情報ｒＡ1と発音期間ｒＡ2（例えば始点および終点）とを指定する。合成音情報ＤＳの単位情報ＵＳが第１合成対象音の発音期間ｓ3を拍点基準で指定するのに対し、合成音情報ＤＲの単位情報ＵＲａは、第２合成対象音の発音期間ｒＡ2を実時間基準で（例えば音響信号Ａの始点から経過した実時間）で指定する。

【００４０】

特徴量情報ＲＢは、第２合成対象音の特徴量（音高）の時間変化を指定する情報であり、特徴量画像６４Tの相異なる編集点αに対応する複数の単位情報ＵＲｂの時系列で構成さ

れる。各単位情報URbは、編集点αの時刻rB1とその編集点αに対応する特徴量の数値rB2とを指定する。時刻rB1は、単位情報URaの発音期間rA2と同様に実時間基準で指定される。すなわち、合成音情報DRは、第２合成対象音の音素と特徴量とを実時間基準で（すなわち実時間を単位として）時系列に指定する。

【００４１】

第２実施形態の制御処理部２２は、対象音編集画面６０に対する利用者からの指示に応じて合成音情報DRを随時に更新する。具体的には、制御処理部２２は、音素列画像６２Tに対する音素の追加や変更が利用者から指示された場合には音素列情報RAの単位情報URaを追加または変更し、特徴量画像６４Tに対する編集点αの追加や変更（移動，削除）が利用者から指示された場合には特徴量情報RBの単位情報URbを追加または変更する。また、利用者は、以上の例示のように所望の音素や編集点αを直接的に指定するほか、拍点基準で指定された既存の第１合成対象音から実時間基準の第２合成対象音への変換を指示することも可能である。利用者から変換が指示されると、制御処理部２２は、拍点基準の合成音情報DSを実時間基準の合成音情報DRに変換して記憶装置１２に格納する。変換後の合成音情報DRについては、利用者が直接的に指定した合成音情報DRと同様に音素列情報RAや特徴量情報RBの編集が可能である。

【００４２】

図７に示すように、第２実施形態の音声合成部２４は、合成音情報DRに応じて音響信号ARを生成する。具体的には、音素列情報RAの各単位情報URaが指定する識別情報rA1に対応する素片情報を音声素片群QAから順次に取得し、各素片情報を、当該単位情報URaの発音期間rA2と特徴量情報RBの各単位情報URbの特徴量rB2（音高）とに調整したうえで相互に連結することで音響信号ARを生成する。そして、音声合成部２４は、合成音情報DSに応じた音響信号ASと波形情報DWが示す音響信号AWと合成音情報DRから生成した音響信号ARとを合成することで音響信号Aを生成する。

【００４３】

図８は、第２実施形態の編集画面３０の模式図である。図８に示すように、第２実施形態の編集画面３０は、第１実施形態と同様の要素（第１領域３１，第２領域３２，第３領域３３）に第４領域３４を追加した構造である。表示制御部２６は、音素列画像６２Bと特徴量画像６４Bとを第４領域３４に配置する。音素列画像６２Bおよび特徴量画像６４Bは、図６の対象音編集画面６０にて実時間軸Xtのもとで表示される音素列画像６２Tおよび特徴量画像６４Tを拍点軸Xb上に展開した画像である。すなわち、音素列画像６２Bは、音素列情報RAが指定する各音素の時系列を拍点軸Xbのもとで表示し、特徴量画像６４Bは、特徴量情報RBが指定する特徴量の時間変化を拍点軸Xbのもとで表示する。表示制御部２６は、対象音編集画面６０内の音素列画像６２Tや特徴量画像６４Tの編集（音素列情報RAや特徴量情報RBの更新）に応じて編集画面３０内の音素列画像６２Bや特徴量画像６４Bを随時に更新する。

【００４４】

第１実施形態と同様に、利用者が速度調整区間４６を指定してテンポt2の変更を指示すると、表示制御部２６は、図９の操作画面５０を表示装置１６に表示させる。第２実施形態の操作画面５０は、第１実施形態と同様のテンポt2の指定領域５２に加えて、第２合成対象音の再生速度V2をテンポt2に連動して変化させるか否か（再生速度V2の変更の許可／禁止）を利用者が選択する操作領域（チェックボックス）５４を含んで構成される。

【００４５】

再生速度V2の変更の禁止が選択された状態（例えば操作領域５４のチェックオフ）では、合成音情報DRが実時間基準で規定する第２合成対象音の再生速度V2が変化しないように、制御処理部２２は、利用者からの指示に応じてテンポt2を変化させた場合でも合成音情報DR（すなわち実時間軸Xt上の第２合成対象音の時系列）を変更しない。したがって、図１０に矢印A1_OFFで示すように、対象音編集画面６０に実時間基準で表示された音素列画像６２Tおよび特徴量画像６４Tはテンポt2の変更の前後で変化しない。他方

、拍点区間δの実時間長はテンポｔ2に応じて変化するから、表示制御部２６は、編集画面３０の第４領域３４に拍点基準で表示される音素列画像６２Bおよび特徴量画像６４Bをテンポｔ2に応じて伸縮する。具体的には、表示制御部２６は、図１１に矢印Ａ2_OFFで示すように、速度調整区間４６内のテンポｔ2が上昇する（拍点区間δの実時間長が短くなる）ほど、編集画面３０内の速度調整区間４６内の音素列画像６２Bおよび特徴量画像６４Bを拍点軸Ｘbの方向に伸長させる。また、速度調整区間４６内のテンポｔ2が低下した場合（拍点区間δの実時間長が長くなる場合）、表示制御部２６は、編集画面３０内の音素列画像６２Bおよび特徴量画像６４Bを拍点軸Ｘbの方向に縮小させる。

【００４６】

他方、再生速度Ｖ2の変更の許可が選択された状態（例えば操作領域５４のチェックオン）では、合成音情報ＤRが実時間基準で規定する第２合成対象音の再生速度Ｖ2がテンポｔ2に連動して変化するように表示制御部２６および制御処理部２２が動作する。すなわち、例えば速度調整区間４６内のテンポｔ2が上昇した場合（拍点区間δの実時間長が短くなる場合）、表示制御部２６は、図１０に矢印Ａ1_ONで示すように、対象音編集画面６０に実時間基準で表示された音素列画像６２Tおよび特徴量画像６４Tを実時間軸Ｘｔの方向に縮小し、制御処理部２２は、第２合成対象音の再生速度Ｖ2がテンポｔ2に応じて上昇するように合成音情報ＤRを更新する。例えば、制御処理部２２は、合成音情報ＤRが指定する音素の発音期間ｒA2や編集点αの時刻ｒB1を実時間軸Ｘｔの前方（遡及方向）に移動させる。また、テンポｔ2が低下した場合、対象音編集画面６０の音素列画像６２Tや特徴量画像６４Tが実時間軸Ｘｔの方向に伸長するように、表示制御部２６による対象音編集画面６０の更新と制御処理部２２による合成音情報ＤRの更新とが実行される。他方、編集画面３０の音素列画像６２Bおよび特徴量画像６４Bは拍点軸Ｘbのもとで表示されるから、図１１の矢印Ａ2_ONで示すように、対象音編集画面６０の音素列画像６２Tおよび特徴量画像６４Tはテンポｔ2の変更の前後で変化しない。

【００４７】

例えば楽曲中の台詞のように拍点に同期させる必要がない音声が第２合成対象音として指定された状態では、拍点のテンポｔ2が変更された場合でも第２合成対象音の再生速度Ｖ2は維持される必要がある。以上のような場合に、利用者は、操作領域５４に対する指示で再生速度Ｖ2の変更を禁止することが可能である。他方、第１合成対象音の変換で第２合成対象音を指定した場合のように第２合成対象音が拍点に同期するように指定された場合、第２合成対象音の再生速度Ｖ2を第１合成対象音とともにテンポｔ2に応じて変化させる必要がある。以上のような場合に、利用者は、操作領域５４に対する指示で再生速度Ｖ2の変更を許可することが可能である。すなわち、第２実施形態によれば、実時間基準で規定された第２合成対象音の再生速度Ｖ2をテンポｔ2に連動して変化させるか否かを利用者が任意に選択できるという利点がある。

【００４８】

＜Ｃ：第３実施形態＞

第２実施形態では、実時間基準で規定される第２合成対象音の再生速度Ｖ2の変更の許可／禁止を制御した。第３実施形態では、拍点基準で規定される第１合成対象音の再生速度Ｖ1の変更の許可／禁止を制御する。

【００４９】

第１実施形態と同様に、利用者がテンポｔ2の変更を指示すると、表示制御部２６は、図１２の操作画面５０を表示装置１６に表示させる。第３実施形態の操作画面５０は、第１実施形態と同様のテンポｔ2の指定領域５２に加えて、第１合成対象音の再生速度Ｖ1をテンポｔ2に連動して変化させるか否か（再生速度Ｖ1の変更の許可／禁止）を利用者が選択する操作領域５６を含んで構成される。

【００５０】

再生速度Ｖ1の変更の禁止が選択された状態では、合成音情報ＤSが拍点基準で規定する第１合成対象音の再生速度Ｖ1が変化しないように表示制御部２６および制御処理部２２が動作する。すなわち、例えば速度調整区間４６内のテンポｔ2が上昇した場合（拍点区

間δの実時間長が短くなる場合）、表示制御部２６は、図１３に矢印ｂ1で示すように、第１領域３１に表示された対象音画像４２（音指示子４２２の時系列）を拍点軸Ｘbの方向に伸長し、制御処理部２２は、第１合成対象音の再生速度Ｖ1がテンポｔ2の変更の前後で維持されるように合成音情報ＤＳを更新する。例えば、制御処理部２２は、合成音情報ＤＳが指定する第１合成対象音の発音期間ｓ3を拍点軸Ｘbの後方（経過方向）に移動させる。また、テンポｔ2が低下した場合（拍点区間δの実時間長が長くなる場合）、図１３に矢印ｂ2で示すように、第１領域３１に拍点基準で表示された対象音画像４２が拍点軸Ｘbの方向に縮小するように、表示制御部２６による対象音画像４２の更新と制御処理部２２による合成音情報ＤＲの更新とが実行される。

【００５１】

他方、再生速度Ｖ2の変更の許可が選択された状態では、表示制御部２６による対象音画像４２の更新や制御処理部２２による合成音情報ＤＲの更新は実行されない。したがって、第１合成対象音の再生速度Ｖ1はテンポｔ2に連動して変化する。すなわち、テンポｔ2が上昇するほど再生速度Ｖ1も上昇する。

【００５２】

例えば楽曲の歌唱音が拍点に同期するように第１合成対象音として指定された状態では、拍点のテンポｔ2に応じて第１合成対象音の再生速度Ｖ1を変化させる必要がある。以上のような場合に、利用者は、操作領域５６に対する指示で再生速度Ｖ1の変更を許可することが可能である。他方、利用者は、楽曲中の台詞のように拍点に同期させる必要がない音声を第１合成対象音として直接的に指定することも可能であり、この場合には、拍点のテンポｔ2が変更された場合でも第１合成対象音の再生速度Ｖ1は維持される必要がある。以上のような場合に、利用者は、操作領域５６に対する指示で再生速度Ｖ1の変更を禁止することが可能である。すなわち、第３実施形態によれば、拍点基準で規定された第１合成対象音の再生速度Ｖ1をテンポｔ2に連動して変化させるか否かを利用者が任意に選択できるという利点がある。

【００５３】

なお、以上の形態では第１実施形態を基礎とした構成を例示したが、再生速度Ｖ1の変更の許可／禁止を制御する構成を第２実施形態に追加した形態も当然に採用され得る。具体的には、拍点基準で規定された第１合成対象音の再生速度Ｖ1の変更の許可／禁止（第３実施形態）と、実時間基準で規定された第２合成対象音の再生速度Ｖ2の変更の許可／禁止（第２実施形態）とが、操作画面５０に対する利用者からの指示に応じて個別に選択される。すなわち、相異なる合成対象音に対応する複数の再生速度の各々について変更の許可／禁止を個別に指定する構成も好適である。例えば、トラック1からトラック3の３種類が併存する場合に、トラック1およびトラック3について再生速度の変更の許可を指示するとともにトラック2について再生速度の変更の禁止を指示することが可能である。

【００５４】

＜Ｄ：第４実施形態＞

第１実施形態では、例えば波形画像４４の各拍点指示子４４２と拍点軸Ｘb上の各基準指示子４０とが拍点軸Ｘb上で同期するように利用者が入力装置１４の操作でテンポｔ2を調整した。第４実施形態では、波形画像４４の各拍点指示子４４２が拍点軸Ｘb上の各基準指示子４０に合致する（クオンタイズされる）ようにテンポｔ2が自動的に（すなわち利用者による調整を必要とせずに）設定される。

【００５５】

利用者が第３領域３３内の速度調整区間４６を選択してテンポｔ2の変更を指示すると、表示制御部２６は、図１４の操作画面５０を表示装置１６に表示させる。第４実施形態の操作画面５０は、第１実施形態と同様のテンポｔ2の指定領域５２に加えて、テンポｔ2の自動的な調整を指示する操作子（「Quantize」コマンドボタン）５８を含んで構成される。利用者は、第１実施形態と同様に指定領域５２に対する操作でテンポｔ2を任意に調整するほか、操作子５８に対する操作でテンポｔ2の自動的な調整を指示することが可能である。

【００５６】
利用者がテンポｔ2の自動的な調整を指示すると、波形画像４４で表現される音響信号ＡＷの各推定拍点（拍点指示子４４２）が拍点軸Ｘb上の各拍点（基準指示子４０）に合致するように、表示制御部２６が速度調整区間４６内の波形画像４４を拍点軸Ｘbの方向に伸縮するとともに制御処理部２２がテンポｔ2を設定する。具体的には、速度調整区間４６内の波形画像４４における特定の推定拍点（例えば速度調整区間４６内の最初の推定拍点）が、拍点軸Ｘb上で当該推定拍点に最も近い拍点に合致するように表示制御部２６が波形画像４４を伸縮し、速度調整区間４６内のテンポｔ2を制御処理部２２が波形画像４４の伸縮の度合に応じた数値に更新する。例えば、表示制御部２６が波形画像４４を伸長させるほどテンポｔ2が低下するように制御処理部２２はテンポｔ2を調整する。

【００５７】
第４実施形態においては、音響信号ＡＷの推定拍点が拍点軸Ｘb上の各拍点に合致するようにテンポｔ2が自動的に調整されるから、利用者が編集画面３０を確認しながらテンポｔ2を調整する第１実施形態と比較して、合成音情報ＤＳに応じた音響信号ＡＳの拍点と波形情報ＤＷに応じた音響信号ＡＷの拍点とを容易に同期させることが可能である。また、テンポｔ2の自動的な調整に加えて、利用者によるテンポｔ2の指定も可能であるから、例えば、拍点指示子４２２が基準指示子４０に充分に接近するように利用者がテンポｔ2を調整してから操作子５８を操作して拍点指示子４２２と基準指示子４０とを合致させるといった運用も可能である。なお、以上の説明では第１実施形態を基礎としたが、テンポｔ2を自動的に調整する第４実施形態の構成は第２実施形態や第３実施形態にも同様に適用され得る。

【００５８】
＜Ｅ：変形例＞
以上の形態には様々な変形が加えられる。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以上の態様は併合され得る。

【００５９】
（１）変形例１
合成音情報ＤＳや波形情報ＤＷや合成音情報ＤＲの内容は適宜に変更される。例えば、合成音情報ＤＳが合成対象音の音素の時系列や特徴量（例えば音高）の時間変化を拍点基準で指定する構成や、合成音情報ＤＲが合成対象音の音高や発音期間を実時間基準で指定する構成も採用され得る。また、例えば音響信号ＡＳや音響信号ＡＲに付加される音響効果の度合の時間変化を波形情報ＤＷが実時間基準で規定する構成も採用され得る。

【００６０】
以上の説明から理解されるように、前述の各形態における合成音情報ＤＳは、第１音楽情報（第１合成対象音の音高ｓ1や発音文字ｓ2）を拍点単位で時系列に指定する第１時系列情報として包括される。他方、前述の各形態における波形情報ＤＷや第２実施形態の合成音情報ＤＲは、第２音楽情報を実時間単位で時系列に指定する第２時系列情報として包括される。第２音楽情報は、波形情報ＤＷが指定する音響信号ＡＷの各振幅値や、合成音情報ＤＲが指定する第２合成対象音の情報（音素の識別情報ｒA1，発音期間ｒA2，特徴量ｒB2）を包含する概念である。

【００６１】
（２）変形例２
以上の各形態では、合成音情報ＤＳと波形情報ＤＷと合成音情報ＤＲとを単一の記憶装置１２に格納したが、合成音情報ＤＳと波形情報ＤＷと合成音情報ＤＲとを別個の記憶装置に格納した構成も採用される。すなわち、拍点基準の時系列情報（合成音情報ＤＳ）を記憶する要素（第１記憶手段）と実時間基準の時系列情報（波形情報ＤＷ，合成音情報ＤＲ）を記憶する要素（第２記憶手段）との別体／一体は本発明において不問である。また、第２実施形態や第３実施形態では波形情報ＤＷは省略され得る。

【００６２】
（３）変形例３

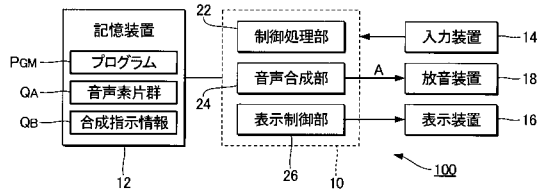
以上の各形態では、音声合成部２４を含む音声合成装置１００を例示したが、音声合成部２４は省略され得る。例えば、表示制御部２６が音楽情報を表示装置１６に表示させる音楽情報表示装置としても本発明は実現され得る。音楽情報表示装置が表示する音楽情報を参照して生成された合成指示情報ＱＢを別体の音声合成装置（音声合成部２４）に提供することで音響信号Ａが生成される。

【符号の説明】

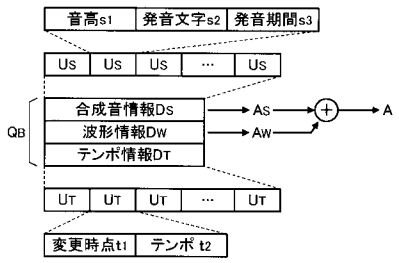
【００６３】

１００……音声合成装置、１０……演算処理装置、１２……記憶装置、１４……入力装置、１６……表示装置、１８……放音装置、２２……制御処理部、２４……音声合成部、２６……表示制御部、３０……編集画面、３１……第１領域、３２……第２領域、３３……第３領域、３４……第４領域、４０……基準指示子、４２……対象音画像、４２２……音指示子、４４……波形画像、４４２……拍点指示子、４６……速度調整区間、５０……操作画面、５２……指定領域、５４……操作領域、５６……操作領域、５８……操作子、６０……対象音編集画面、６２Ｔ，６２Ｂ……音素列画像６２Ｔ、６４Ｔ，６４Ｂ……特徴量画像、α……編集点、Ｘｂ……拍点軸、Ｘｔ……実時間軸、Ｘｐ……音高軸、

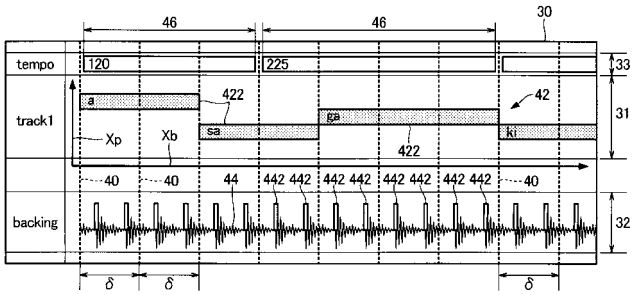
【図１】



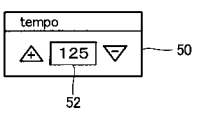
【図２】



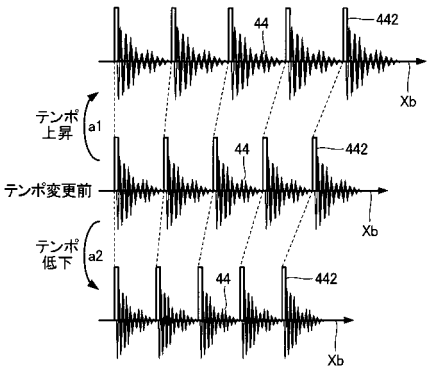
【図３】



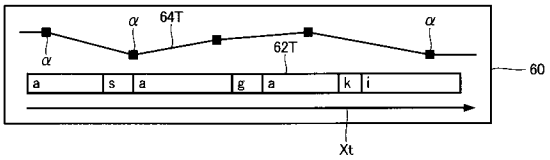
【図４】



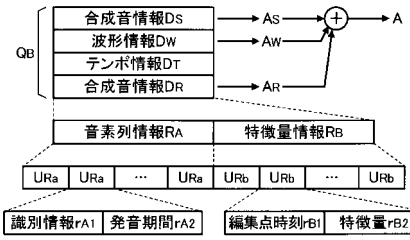
【図５】



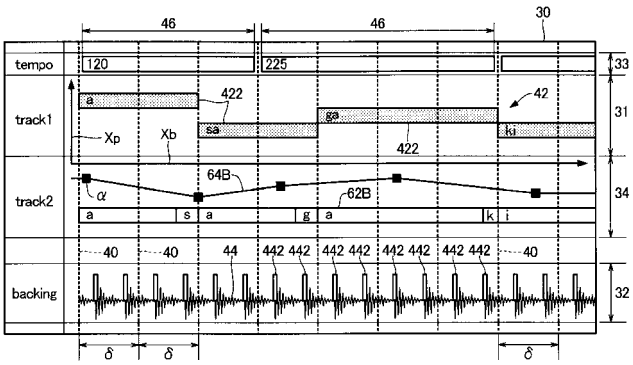
【図６】



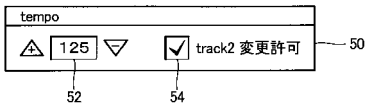
【図７】



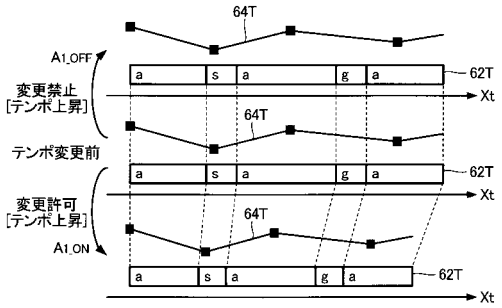
【図８】



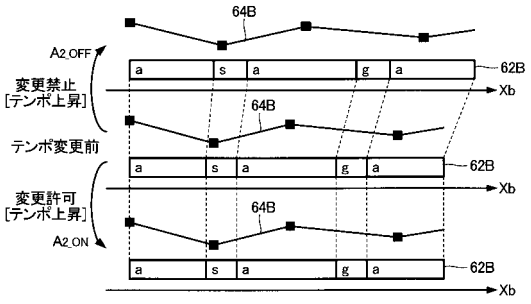
【図９】



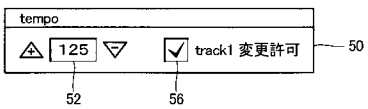
【図１０】



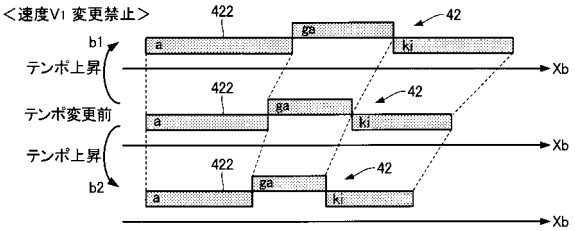
【図１１】



【図１２】



【図１３】



【図１４】

