

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7366282号
(P7366282)

(45)発行日 令和5年10月20日(2023.10.20)

(24)登録日 令和5年10月12日(2023.10.12)

(51)国際特許分類		F I	
G 1 0 H	1/00 (2006.01)	G 1 0 H	1/00 1 0 2 Z
G 1 0 L	25/51 (2013.01)	G 1 0 L	25/51 3 0 0
G 1 0 H	1/36 (2006.01)	G 1 0 H	1/36
G 1 0 G	1/00 (2006.01)	G 1 0 G	1/00

請求項の数 12 (全14頁)

(21)出願番号	特願2022-550690(P2022-550690)	(73)特許権者	522331622
(86)(22)出願日	令和3年2月1日(2021.2.1)		アンテスコフォ
(65)公表番号	特表2023-515122(P2023-515122 A)		フランス・7 5 0 0 3・パリ・リュ・メ レ・3 8
(43)公表日	令和5年4月12日(2023.4.12)	(74)代理人	100108453
(86)国際出願番号	PCT/EP2021/052250		弁理士 村山 靖彦
(87)国際公開番号	WO2021/165023	(74)代理人	100110364
(87)国際公開日	令和3年8月26日(2021.8.26)		弁理士 実広 信哉
審査請求日	令和4年10月18日(2022.10.18)	(74)代理人	100133400
(31)優先権主張番号	20305168.5		弁理士 阿部 達彦
(32)優先日	令和2年2月20日(2020.2.20)	(72)発明者	ジョセ・エチェベステ
(33)優先権主張国・地域又は機関	欧州特許庁(EP)		フランス・7 5 0 1 1・パリ・リュ・ボ バンクール・3 2
		(72)発明者	フィリップ・キュヴィリエ
			フランス・7 5 0 0 5・パリ・プールヴ 最終頁に続く

(54)【発明の名称】 ユーザの楽曲演奏時のあらかじめ記録された楽曲伴奏の改善された同期

(57)【特許請求の範囲】

【請求項 1】

あらかじめ記録された楽曲伴奏をユーザの楽曲演奏に同期させるための方法であって、
前記ユーザの楽曲演奏が、処理ユニットに供給する入力音響信号を届ける少なくとも1
つのマイクロフォンによって取り込まれ、

前記処理ユニットが、前記あらかじめ記録された楽曲伴奏のデータを記憶するためのメ
モリを備え、前記ユーザのために前記楽曲伴奏を演奏する少なくとも1つのラウドスピー
カーに供給するために、前記あらかじめ記録された楽曲伴奏データに基づいて出力音響信
号を提供し、

前記処理ユニットが、

前記入力音響信号中の音楽イベントを検出し、前記ユーザの楽曲演奏におけるテンポ
を決定するために、前記入力音響信号を分析し、

少なくとも、前記検出された音楽イベントのタイミングと、前記演奏された楽曲伴奏
の音楽イベントのタイミングとの間のラグdiffを決定するために、前記検出された音楽イ
ベントを前記あらかじめ記録された楽曲伴奏データと比較し、前記ラグdiffが補正され、

前記出力音響信号のタイミングを、

* 前記ラグdiffと、

*

【数 1】

$$F(x) = \begin{cases} \frac{x^2}{w^2} + (\text{\$tempo} - \frac{2}{w})x + 1 & \text{if } \text{\$diff} > 0 \\ -\frac{x^2}{w^2} + (\text{\$tempo} + \frac{2}{w})x - 1 & \text{if } \text{\$diff} < 0 \end{cases}$$

によって与えられる同期関数Fと
に基づいて適合させ、

ここでxは時間変数であり、\$tempoは前記ユーザの楽曲演奏における前記決定されたテンポであり、wは前記ラグdiffの補正の継続時間である、方法。

10

【請求項2】

前記楽曲伴奏データが、楽曲スコアを定義し、変数xが、前記楽曲スコアの可変の拍数の継続時間に対応する時間的値である、請求項1に記載の方法。

【請求項3】

wが、前記楽曲伴奏データによって定義される楽曲スコア上で少なくとも1拍の継続時間を有する、請求項1または2に記載の方法。

【請求項4】

前記継続時間wが選択される、請求項1、2、および3のいずれか一項に記載の方法。

【請求項5】

20

前記伴奏データが、楽曲スコアを定義し、前記スコア上の演奏者の演奏の位置posが、 $\text{pos}(x) = \text{\$tempo} \cdot x$ として定義される線形関係によって予想され、ここでxは前記楽曲スコア上でカウントされる音楽の拍の数であり、ラグdiffが検出される場合、前記ラグ時間diffに対応する拍数 x_{diff} を定義するために、前記同期関数 $F(x)$ が使用され、

$$F(x_{\text{diff}}) - \text{pos}(x_{\text{diff}}) = \text{diff}$$

ようになる、請求項1から4のいずれか一項に記載の方法。

【請求項6】

予測が、

【数2】

30

$$A(t) \stackrel{\text{def}}{=} F(t - t_0 + x_{\text{diff}}) + p$$

によって与えられる変換関数A(t)を適用することによって、次の拍 $x_{\text{diff}} + w$ まで、前記同期関数F(x)に基づいて決定され、

ここでpは、現在の時間 t_0 における前記楽曲スコア上の前記演奏者の演奏の現在の位置である、請求項5に記載の方法。

【請求項7】

40

前記伴奏データが楽曲スコアを定義し、前記処理ユニットが、将来の同期時間 t_{sync} における前記楽曲スコア上の演奏者の演奏の将来の位置をさらに推定し、前記将来の同期時間 t_{sync} まで前記出力音響信号に適用するために前記楽曲伴奏のテンポ(e2)を決定する、請求項1から6のいずれか一項に記載の方法。

【請求項8】

ctempoと示す、前記出力音響信号に適用する前記楽曲伴奏の前記テンポが、現在の時間 t_0 におけるA(t)の導関数として決定され、

$$\text{ctempo} = A'(t_0) = F'(x_{\text{diff}})$$

ようになる、請求項6と組み合わせて行われる、請求項7に記載の方法。

【請求項9】

50

前記入力音響信号中の前記音楽イベントの前記決定が、
前記入力音響信号から音響特徴を抽出するステップと、
少なくとも前記伴奏中の音楽イベントを決定するために、前記あらかじめ記録された
楽曲伴奏の前記記憶されたデータを使用するステップと、
前記記憶されたデータから決定された前記音楽イベントに基づいて、前記入力音響特
徴に音楽イベントを割り当てるステップと
を含む、請求項1から8のいずれか一項に記載の方法。

【請求項10】

あらかじめ記録された楽曲伴奏をユーザの楽曲演奏に同期させるためのデバイスであっ
て、請求項1から9のいずれか一項に記載の方法を行うために処理ユニットを備える、デバ
イス。

10

【請求項11】

命令を含むコンピュータプログラムであって、前記プログラムが処理ユニットによって
実行されると、前記命令が前記処理ユニットに請求項1から9のいずれか一項に記載の方法
を実行させる、コンピュータプログラム。

【請求項12】

命令を含むコンピュータ可読媒体であって、処理ユニットによって実行されると、前記
命令が前記コンピュータに請求項1から9のいずれか一項に記載の方法を実行させる、コン
ピュータ可読媒体。

【発明の詳細な説明】

20

【技術分野】

【0001】

本開示は、人間の演奏者と、その人間の演奏者の伴奏を提供するあらかじめ記録された
音楽データとのリアルタイムの音楽同期性を可能にするデータ処理に関する。

【背景技術】

【0002】

許容できる音楽行動を実現するために、演奏者の音楽意図を把握し、それらをあらかじ
め記録された伴奏の音楽意図にマップすることが目的である。

【0003】

いくつかの知られているシステムは、演奏者と伴奏との間のリアルタイムの音楽同期性
の課題に対処している。

30

【0004】

文献D1:Christopher Raphael(2010)、「Music Plus One and Machine Learning
」、Proceedings of the 27th International Conference on Machine Learning (I
CML)、Haifa、Israel、21～28頁は、演奏者の意図が、同じ演奏者の実際の演奏につい
て訓練されるモデルから予測される学習システムに関するものである。訓練に対するデー
タ利用可能性の問題にもかかわらず、同期はここでは、イベントの確率的パラメータでは
なく、ハイレベルな音楽パラメータ(音楽学的データなど)に左右される。その上、統計的
または確率的予測は、セッション間の(および同じ演奏者についての)演奏の極端な変動性
を弱体化させる。さらに、この手法は、演奏者イベントをコンピュータアクションと同期
させることに依存する。コンピュータアクションは、ハイレベルな音楽パラメータをモデ
ル化せず、したがって非実用的である。

40

【0005】

文献D2:Roger B Dannenberg(1997)、「Abstract time warping of compound e
vents and signals」、Computer Music Journal、61～70頁は、演奏者のテンポが連
続的であり、2つのイベント間で保たれるという基本的仮定を採用し、楽曲位置の区分的
線形予測が同期に使用されることになる。どんな現実世界のセットアップでも、テンポ不
連続性は、そのような近似の失敗につながることである。さらに、この手法は、演奏者の
時間マップしか考慮に入れず、あらかじめ記録された伴奏時間マップ(固定されていると仮
定する)を弱体化させ、したがって重要なハイレベルな音楽知識を見落とす。

50

【 0 0 0 6 】

文献D3:Arshia Cont、Jose Echeveste、Jean-Louis Giavitto、およびFlorent Jacquemard.(2012)、「Correct Automatic Accompaniment Despite Machine Listening or Human Errors in Antescofo」、Proceedings of International Computer Music Conference (ICMC), Ljubljana (Slovenia)は、演奏者の時間マップを推定するために脳の認知モデルを用いる予想の概念を組み入れている。伴奏同期に対してハイレベルな音楽知識を組み入れるために、2つのタイプの同期が導入される。いくつかの重要な位置が厳密に同期されることを確実にするために、タイト同期(Tight Synchronization)が使用される。

適切ながら、それらの解決策は、結果として得られる時間マップに不連続性をもたらす。連続したオーディオまたはビデオストリームを同期するとき、そのような不連続性は、回避されなければならない。スムーズ同期(Smooth Synchronization)は、結果として得られる伴奏のテンポが演奏者のテンポに等しいと仮定し、その値を使用してその位置を予測することによって、結果として得られる連続的な時間マップを生成しようと試みる。

この適切なテンポ検出にもかかわらず、リアルタイムのテンポはエラーになりやすく、予測できない不連続性につながることもある。さらに、同じセッションにおける2つの戦略の共存が、結果として得られる時間マップにさらなる不連続性を生む。

【 0 0 0 7 】

文献D4:Dawen Liang、Guangyu Xia、およびRoger B Dannenberg(2011)、「A framework for coordination and synchronization of media」、Proceedings of the International Conference on New Interfaces for Musical Expression(167 ~ 172頁)は、基準伴奏時間マップへの値の収束を目標として、時間マップを動的に同期させるために、上記のタイトなどの散発的同期と、ルーズ(Loose)などのテンポのみの同期との間の妥協点を提案する。伴奏が時間 t における偏差を補正するようにし、それが $t+w$ において収束するように、未来への音楽継続時間 w にわたる一定ウィンドウ(constant window)が使用される。これは、音楽位置出力で区分的線形である連続曲線をもたらす。

この戦略はしかしながら、2つの欠点を有する。

- 連続的位置にもかかわらず、テンポ不連続性が依然として存在する。演奏者のテンポが変わらないとき、伴奏テンポが変わることがあるので、そのような不連続性が、演奏者への誤ったフィードバックにつながる。
- 一定ウィンドウ処理は、中間更新と一致しない。一例は、予測される演奏者の時間マップを変えることはない時間 t における初期ラグの存在が、永続的なラグにつながることである。

【 先行技術文献 】

【 非特許文献 】

【 0 0 0 8 】

【 文献 】Christopher Raphael(2010)、「Music Plus One and Machine Learning」、Proceedings of the 27th International Conference on Machine Learning (ICML)、Haifa、Israel、21 ~ 28頁

Roger B Dannenberg(1997)、「Abstract time warping of compound events and signals」、Computer Music Journal、61 ~ 70頁

Arshia Cont、Jose Echeveste、Jean-Louis Giavitto、およびFlorent Jacquemard.(2012)、「Correct Automatic Accompaniment Despite Machine Listening or Human Errors in Antescofo」、Proceedings of International Computer Music Conference (ICMC), Ljubljana (Slovenia)

Dawen Liang、Guangyu Xia、およびRoger B Dannenberg(2011)、「A framework for coordination and synchronization of media」、Proceedings of the International Conference on New Interfaces for Musical Expression(167 ~ 172頁)

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 9 】

本開示は、この状況を改善することを目的とする。

【課題を解決するための手段】

【 0 0 1 0 】

そのために、あらかじめ記録された楽曲伴奏をユーザの楽曲演奏に同期させるための方法が提案され、前記ユーザの楽曲演奏が、処理ユニットに供給する入力音響信号を届ける少なくとも1つのマイクロフォンによって取り込まれ、前記処理ユニットが、あらかじめ記録された楽曲伴奏のデータを記憶するためのメモリを備え、前記ユーザのために楽曲伴奏を演奏する少なくとも1つのラウドスピーカーに供給するために、前記あらかじめ記録された楽曲伴奏データに基づいて出力音響信号を提供し、前記処理ユニットが、

- 入力音響信号中の音楽イベントを検出し、前記ユーザの楽曲演奏におけるテンポを決定するために、入力音響信号を分析し、
- 少なくとも、検出された音楽イベントのタイミングと、演奏された楽曲伴奏の音楽イベントのタイミングとの間のラグdiffを決定するために、検出された音楽イベントをあらかじめ記録された楽曲伴奏データと比較し、前記ラグdiffが補正され、
- 出力音響信号のタイミングを、
 - * 前記ラグdiffと、
 - *

10

【 0 0 1 1 】

【数 1】

20

$$F(x) = \begin{cases} \frac{x^2}{w^2} + (\text{\$tempo} - \frac{2}{w})x + 1 & \text{if } diff > 0 \\ -\frac{x^2}{w^2} + (\text{\$tempo} + \frac{2}{w})x - 1 & \text{if } diff < 0 \end{cases}$$

【 0 0 1 2 】

によって与えられる同期関数Fとに基づいて適合させ、ここでxは時間変数であり、\$tempoはユーザの楽曲演奏における決定されたテンポであり、wは前記ラグdiffの補正の継続時間である。

30

【 0 0 1 3 】

「時間マップ」という概念は、したがって、あらかじめ記録された伴奏と比較して人間の演奏者から入ってくる音楽意図をモデル化するために使用され得る。時間マップは、物理的時間tを音楽的時間p(拍単位)にマップする関数である。

非リアルタイム(またはオフライン)のセットアップにおいて、デバイスからのテンポ推定が正しいという強い仮定を与えられるならば、時間マップ位置pは、拍子に時間0からtのこのテンポを掛けた整数である。

しかしながら、演奏者が楽曲スコアに設定されているテンポに従わないとき、伴奏の現在の演奏における推定テンポは、補正継続時間wによって定義される近い将来に適合される必要があり、同期関数Fの使用は、その補正継続時間の後に現在のユーザのテンポへと収束に達することを確実にする。

40

【 0 0 1 4 】

一実施形態では、前記楽曲伴奏データは、楽曲スコアを定義し、ここで変数xは、前記楽曲スコアの可変の拍数の継続時間に対応する時間的値である。

【 0 0 1 5 】

一実施形態では、前記補正継続時間wは、前記楽曲伴奏データによって定義される楽曲スコア上で少なくとも1拍の継続時間を有する。

【 0 0 1 6 】

一実施形態では、前記補正継続時間wは、選択される。

50

好ましくはそれは、1拍の継続時間に設定され得るが、場合によっては、たとえば前記処理ユニットの入力を介して入力できるユーザの選択に従って、より多くに設定され得る。

【0017】

伴奏データが楽曲スコアを定義する一実施形態では、前記スコア上の演奏者の演奏の位置 pos は、 $pos(x)=\$tempo \cdot x$ として定義される線形関係によって予想され、ここで x は前記楽曲スコア上でカウントされる楽曲拍の数であり、ラグ $diff$ が検出される場合、同期関数 $F(x)$ は、次いで前記ラグ時間 $diff$ に対応する拍数 x_{diff} を定義するために使用され、 $F(x_{diff}) - pos(x_{diff}) = diff$ のようになる。

【0018】

この実施形態では、予測は、

10

【0019】

【数2】

$$A(t) \stackrel{\text{def}}{=} F(t - t_0 + x_{diff}) + p$$

【0020】

によって与えられる変換関数 $A(t)$ を適用することによって次の拍 $x_{diff} + w$ まで、前記同期関数 $F(x)$ に基づいて決定される。

20

ここで p は、現在の時間 t_0 における楽曲スコア上の演奏者の演奏の現在の位置である。

【0021】

前記伴奏データが楽曲スコアを定義する一実施形態では、処理ユニットが、将来の同期時間 t_{sync} における前記楽曲スコア上の演奏者の演奏の将来の位置をさらに推定し、前記将来の同期時間 t_{sync} まで出力音響信号に適用するために、楽曲伴奏のテンポ(以下で提示する図3の参照符号e2)を決定する。

【0022】

この実施形態では、変換関数 $A(t)$ が使用されるとき、出力音響信号に適用する楽曲伴奏のテンポは、以下のように、現在の時間 t_0 における $A(t)$ の導関数として決定される。

30

$$tempo = A'(t_0) = F'(x_{diff})$$

(これは分析的にわかっている)。

【0023】

一実施形態では、前記入力音響信号中の前記音楽イベントの決定は、

- 前記入力音響信号から音響特徴(たとえば、時間による音響圧力、または認識された高調波周波数)を抽出することと、
- 少なくとも伴奏中の音楽イベントを決定するために、あらかじめ記録された楽曲伴奏の前記記憶されたデータを使用することと、
- 前記記憶されたデータから決定された音楽イベントに基づいて、前記入力音響特徴に音楽イベント(たとえば特定の音符のアタックタイム)を割り当てることと

40

を含む。
事実上、音楽イベントの割当ては、楽曲スコア上に、たとえばソロパート上で行われ、したがってこれによって、「伴奏」自体によってではなく、決定され得る。これらは、一般的にはMIDIなどの記号音楽表記形式とすることができる。したがって、「伴奏楽曲スコアの記憶されたデータ」という表現は、広く解釈されるものとし、そのようなデータが、伴奏自体ではないソロトラックの楽曲スコアをさらに含む状況を包含し得る。

楽曲スコアイベントの関連付けは、より一般的にはあらかじめ記録された伴奏(時間マップ)において行われる。

【0024】

本開示はまた、上記で提示した方法を行うための処理ユニットを備える、あらかじめ記

50

録された楽曲伴奏をユーザの楽曲演奏に同期させるためのデバイス为目标とする。

【0025】

本開示はまた、命令を含むコンピュータプログラムを目標とし、プログラムが処理ユニットによって実行されると、命令は、処理ユニットに上記で提示した方法を実行させる。

【0026】

本開示はまた、命令を含むコンピュータ可読媒体を目標とし、処理ユニットによって実行されると、命令は、コンピュータに方法を実行させる。

【0027】

したがって、演奏者とあらかじめ記録された伴奏との間のリアルタイムの同期性を実現するために、本開示は特に最先端技術における以下の欠点に対処する。

10

- 演奏者の時間マップは、デバイスから入って来るものとして与えられず、あらかじめ記録された伴奏における固有の時間マップなどのハイレベルな音楽知識を考慮に入れて予測される。

- 伴奏出力に対して時間マップを予測するとき、テンポ(および必ずではないが位置)の不連続性は、(演奏者によって)音楽的にも、(オーディオまたはビデオストリームなどの連続的メディアに対して)技術的にも許容できない。これは単独で、区分的線形予測に基づく従来技術の手法すべてを不適格と見なし得る。

- あらかじめ記録された伴奏を駆動するための結果として得られるリアルタイムの時間マップは、演奏者の時間マップ(意図の把握)とあらかじめ記録された伴奏時間マップ(ハイレベルな音楽知識)の両方に依存している。

20

【0028】

実施形態のさらなる詳細および利点については、以下の詳細な説明で提供し、添付の図面に表示する。

【図面の簡単な説明】

【0029】

【図1】前述の方法を実施するデバイスの実施形態の一例を示す図である。

【図2】一実施形態による前述の方法のステップを含むアルゴリズムの一例を示す図である。

【図3a】同期関数 $F(x)$ を使用した同期時間マップおよび対応する演奏者時間マップの一例を示す図である。

30

【図3b】同期関数 $F(x)$ を使用した同期時間マップおよび対応する演奏者時間マップの一例を示す図である。

【発明を実施するための形態】

【0030】

本開示は、あらかじめ記録された伴奏をリアルタイムで演奏者に同期させる問題を解決することを提案する。この目的のために、デバイスDIS(以下で説明する図1の例に示すように)が使用される。

【0031】

デバイスDISは、一実施形態では、少なくとも

- 入力インターフェースINPと、
- ストレージメモリMEM、およびメモリMEMと協働するプロセッサPROCを含む処理ユニットPUと、
- 出力インターフェースOUTと

を備える。

40

【0032】

メモリMEMは、本開示によれば、とりわけ、コンピュータプログラムの命令データを記憶することができる。

【0033】

さらに、楽曲伴奏データが、処理ユニットに(たとえば、メモリMEMに)記憶される。楽曲伴奏データはしたがって、あらかじめ記録された楽曲伴奏データに基づいて出力音響信

50

号を少なくとも1つのラウドスピーカーSPK(バッフルまたはイヤフォン)に供給するために、出力インターフェースOUTを駆動させるように、プロセッサPROCによって読み出される。

【0034】

デバイスDISは、(図1の破線で示すように)独立したハードウェアを含むことができる、または代替的に処理ユニットPUと共有されるハードウェア(すなわち同じプロセッサおよび場合によっては同じメモリユニット)で作られ得る機械リスニングモジュールMLMをさらに備える。

【0035】

ユーザUSは、ラウドスピーカーSPKによって演奏される伴奏楽曲を聞くことができ、伴奏楽曲上の楽器とともに、このようにして入力インターフェースINPに接続されたマイクロフォンMICによって取り込まれた音を発して、演奏することができる。マイクロフォンMICは、ユーザの楽器(エレキギターなど)に組み込まれている、または(音声もしくは音響機器録音のために)分けられていることがある。取り込まれた音データは、次いで機械リスニングモジュールMLMによって、およびより一般的には処理ユニットPUによって処理される。

10

【0036】

より詳細には、取り込まれた音データは、伴奏楽曲と比較され、ユーザによって演奏される楽曲の遅れまたは進みを識別し、伴奏楽曲の演奏の速度をユーザの演奏に適合させるために処理される。たとえば、伴奏楽曲のテンポは、それに応じて適合され得る。伴奏楽曲とユーザによって演奏される楽曲との間の、モジュールMLMによって検出される時間差は、以下では、現在の時間tにおける「ラグ」と呼び、diffと示す。

20

【0037】

より詳細には、演奏者イベントは、機械リスニングモジュールMLMによってリアルタイムで検出可能であり、機械リスニングモジュールMLMはその場合、音楽イベントの連符(t-uplets)およびそのようなイベントのリアルタイム検出に関するテンポデータを楽曲スコアから出力する。この実施形態は、たとえば、Cont(2010)に開示されるものと同様とすることができる。機械リスニングモジュールMLMが処理ユニットPUから切り離されたハードウェアを有する実施形態では、モジュールMLMは、したがって交換可能であり、したがって、演奏者の演奏を聴くことによって、リアルタイムで所与の楽曲スコア上に「イベント」、および任意選択で以後テンポを提供する任意のモジュールとすることができる。

30

【0038】

上記のように、機械リスニングモジュールMLMは、好ましくは「リアルタイムで」、理想的には現在の通常のリスニングアルゴリズムの大部分における知覚しきい値(イベントに反応する能力)に対応する15ミリ秒未満のラグで動作する。

【0039】

一方ではあらかじめ記録された伴奏楽曲データのおかげで、および他方では演奏者の演奏でのテンポ認識のおかげで、処理ユニットPUは、動的同期を行う。各リアルタイムのインスタンスtで、処理ユニット(PU)は、前の時間t-εにおけるそれ自体の前の予測、ならびに機械リスニングからの入って来るイベントおよびテンポを入力として受け取る。結果として得られる出力は、時間tにおける予測を含む伴奏時間マップである。

40

【0040】

同期は、動的に計算されるラグ依存ウィンドウ(以下wと示す)に基づく、時間tにおける予測出力のおかげで、動的かつ適応可能である。動的同期戦略が導入され、その値は、後の時間t_syncに収束すると数学的に保証される。同期予想範囲(synchronization anticipation horizon)t_sync自体は、前のインスタンスおよび環境からのフィードバックに関して、時間tにおける計算されたラグ時間に依存している。

【0041】

適応同期戦略の結果は、一貫したものとなる(同じセットアップは同じ同期予測になる)。適応同期戦略はまた、対話型コンテキストに適応する。

50

【 0 0 4 2 】

デバイスDISは、演奏者のイベントおよびテンポをライブ入力として受け取り、あらかじめ記録された伴奏に対する予測を出力し、起動に先立って自由裁量であらかじめ記録された伴奏と楽曲スコアの両方を有する。デバイスDISの役割は、(ライブ入力の結果としての)演奏者の時間マップを使用し、対応する同期時間マップを動的に構築することである。

【 0 0 4 3 】

(最先端技術のように)一定ウィンドウ長に依存するのではなく、パラメータwは、ここでは剛性パラメータとして解釈される。一般に、wは、スコアの固定拍数(たとえば、4/4拍子の四分音符に対応する1拍)に対応することができる。その時間現在値(time current value)tvは、伴奏の実際のテンポで与えられ得る($tv=w*$ 実際のテンポ)が、これは必ずしも現在の演奏者のテンポに対応するとは限らない。予測ウィンドウ長wは、時間tにおける現在のラグdiffの関数として(図3に関して以下で詳述するように)動的に決定され、後の同期時間 t_{sync} まで収束を確実にする。

10

【 0 0 4 4 】

一実施形態では、同期関数Fが導入され、その役割は、同期時間マップの構築を助けること、およびテンポが短い時間枠で、一定の値であると考えられる理想的なセットアップにおいてラグdiffを補正することである。(楽曲スコア上の)演奏者の位置pおよび時間tにおける以下で「\$tempo」と示す演奏者のテンポを仮定すると、Fは、時間マップ点(0,1)を($w, w*\$tempo$)に結びつける二次関数であり、その導関数がパラメータ\$tempoに等しいことを調べる。楽曲スコア上の演奏者のリアルタイムの音楽位置と、同じスコア上の伴奏トラックのそれ(どちらも拍単位)との間の時間tにおけるラグは、diffとして示される。したがって、パラメータdiffは、リアルタイムの検出された演奏者のイベントの拍単位の楽曲スコア上の位置と、同期されることになる伴奏楽曲の楽曲スコア(拍単位)上の位置との間の違いを正確に反映する。

20

【 0 0 4 5 】

ここで、同期関数Fを以下の通りに表すことができることを示す:

【 0 0 4 6 】

【数 3】

$$F(x) = \begin{cases} \frac{x^2}{w^2} + (\$tempo - \frac{2}{w})x + 1 & \text{if } diff > 0 \\ -\frac{x^2}{w^2} + (\$tempo + \frac{2}{w})x - 1 & \text{if } diff < 0 \end{cases}$$

30

【 0 0 4 7 】

diff=0の場合、F(x)は単に、 $F(x)=\$tempo*x$ となり、ここで\$tempoは、モジュールMLMによって提供される実際のテンポ値であり、wは、演奏者の演奏時の楽曲伴奏の次の調整までラグdiffを補正するためにかかった時間に最終的に対応する予測ウィンドウである。

【 0 0 4 8 】

40

時間tにおいて検出された任意のイベント、および拍の進んだ伴奏ラグdiffについて、式 $F(x)-\$tempo*x=diff$ の単一の解 x_{diff} があることをさらに示す。この一意の解は、それに関して予測が計算される適応コンテキストを定義し、以下のように、 x_{diff} から伴奏マップの部分再定義する。

【 0 0 4 9 】

【数 4】

50

$$A(t) \stackrel{\text{def}}{=} F(t - t_0 + x_{\text{diff}}) + p$$

【 0 0 5 0 】

適応関数A(t)の詳細な説明は、以下で行う。

【 0 0 5 1 】

構造によって、伴奏同期の時間マップは、時間 $t_{\text{sync}}=t+w-x_{\text{diff}}$ における位置およびテンポにおいて、演奏者時間マップに収束する。この数学的構造は、同期時間 t_{sync} までテンポの連続性を確実にする。

10

【 0 0 5 2 】

図3は、イベントが検出され、伴奏の初期ラグが拍の進んだdiffである(図3a)、時間tにおいて、伴奏時間マップを更新するための適応動的同期を示す。tからの伴奏マップは、関数Fの置換された部分として定義される。F(x)によって構築される同期時間マップは、図3(a)に示され、演奏者時間マップへのその置換が、図3(b)に示されている。位置およびテンポは、時間 t_{sync} で収束し、演奏者テンポはその区間において一定のままであると仮定する。この時間マップは、システムが人間の演奏者と対話するごとに常に再評価される。時間 t_{sync} までのテンポの連続性に気づくことができる。

20

【 0 0 5 3 】

図3の簡単な説明を、以下の通り行うことができる。前の予測から、演奏者の演奏が有する(拍xでカウントされる)予想位置posが、 $\text{pos}(x)=\text{tempo} \cdot x$ などの線形関係によって決定される。これは、図3aの斜めの破線に対応する。しかしながら、演奏者の演奏の位置pと予想位置posとの間にラグdiffが検出される。同期関数F(x)は、上記で定義したように計算され、 $F(x_{\text{diff}})-\text{pos}(x_{\text{diff}})=\text{diff}$ のように x_{diff} が計算される。したがって次の拍 $x_{\text{diff}}+w$ まで、F(x)に基づいて、予測を決定することができる。これは、図3aの破線の長方形に対応する。図3aのこの「長方形」はむしろ、図3bの演奏者時間マップにインポートされ、

【 0 0 5 4 】

【数5】

30

$$A(t) \stackrel{\text{def}}{=} F(t - t_0 + x_{\text{diff}}) + p$$

【 0 0 5 5 】

によって与えられる変換関数A(t)を適用することによって置換される。

ここでpは、現在の時間 t_0 におけるスコア上の演奏者の演奏の現在の位置である。したがってA(t)は、演奏者の演奏が将来の時間 t_{sync} に有する正しい位置を与えるために計算することができる。少なくともこの同期時間 t_{sync} まで、伴奏のテンポは適合される。それは、前の傾き $e1$ と比較するため、新しい傾き $e2$ (図3bの斜めの破線)に対応する。訂正されたテンポ $ctempo$ は、このように、現在の時間 t_0 におけるA(t)の導関数として与えることができ、すなわち以下の通りである。

40

$$ctempo=A'(t_0)=F'(x_{\text{diff}})$$

これは分析的にわかっている。

【 0 0 5 6 】

次に図2を参照すると、演奏者の演奏に関係する入力信号を受信して、ステップS1が始まる。ステップS2において、あらかじめ記録された楽曲伴奏データにおいて定義されている楽曲スコア中のイベントに関係している演奏者の演奏における音楽イベントを識別する

50

ために、音響特徴が入力信号から抽出される。ステップS3において、最新の検出されたイベントのタイミングが、スコア内の対応するイベントのタイミングと比較され、タイミング差に対応するタイムラグdiffが決定される。

【0057】

その時間ラグおよび選択された継続時間w(一般的には楽曲スコアにおける選択された拍数の継続時間)に基づいて、ステップS4において同期関数 $F(x)$ を決定することができる。次いで、ステップS5において、 x_{diff} が、 $F(x_{diff}) - \text{tempo} * x_{diff} = \text{diff}$ によって与えられる唯一の解とすることができる。

【0058】

x_{diff} の決定は次いで、図3aおよび図3bを参照しながら上記で説明したように同期マップから演奏者時間マップにシフトするために、ステップS6において決定された変換関数 $A(t)$ を使用することを可能にする。演奏者時間マップでは、ステップS7において、あらかじめ記録された伴奏データに基づいて演奏される出力信号のテンポは、図3bに示すように、出力信号の楽曲スコア上の位置を、将来の次の同期時間 t_{sync} での入力信号の位置まで滑らかに調整するために、(図3bの傾き $e1$ から傾き $e2$ に)訂正することができる。ステップS8におけるその同期時間 t_{sync} の後(テストS8からの矢印Y)、プロセスは、入力信号から新しい特徴を抽出することによって再び実施され得る。

【0059】

性質上、この実施形態は、以下の利点に到達することに寄与する。

- ・ この実施形態は、最先端技術における一貫性問題を解決する。この実施形態は、初期ラグに自動的に適合し、コンテキストに基づいてその範囲を適合させる。数学的形式(mathematical formalism)は、解と全単射である。これは、同一の演奏者時間マップが同じ同期軌道に至ることを意味するが、伝統的な一定ウィンドウではこの値は、コンテキストおよびパラメータに基づいて異なる。

- ・ 最先端技術がすべての利用可能な方法において不連続性を示すのに対し、この方法は、時間 t_{sync} におけるテンポ連続性を確実にする。

- ・ 適応戦略は、タイトおよびルーズとして上記で説明した2つの極値間に、単一フレームワーク内で妥協点を提供する。タイト戦略は、剛性パラメータwの低い値に対応するが、ルーズ戦略は、wのより高い値に対応する。

- ・ この戦略は、計算効率が良い。予測時間マップが変わらない限り、伴奏同期は伴奏時間マップを使用して1度だけ計算される。最先端技術は、変化にかかわらず、対話のあらゆる段階で計算および予測を必要とする。

【0060】

さらに、ハイレベルな音楽知識が、時間マップの形で同期機構に統合され得る。この目的で、予測が時間マップ上の非線形曲線に拡張される。この拡張は、アツチェレンディ(accelerendi)およびフェルマータなどの(すなわち適応できるテンポを持つ)音楽表現および演奏者のタイミングの他の一般的な表現の音楽的指定を統合するための形式を可能にする。この追加はまた、既存のデータからそのようなパラメータを自動的に知る可能性を容易にする。

- ・ それは、存在するならば、ハイレベルな音楽知識を、収束の証明とともに数学的形式を使用して既存のフレームワークに追加することを可能にし、通常の従来技術の手作業の(hand-engineering)方法を克服する。

- ・ それは、区分的線形予測につながる通常の従来技術における「一定のテンポ」の近似を、より現実的な非線形テンポの予測に拡張する。

- ・ それは、表現的に活用するために、演奏者またはあらかじめ記録された伴奏のいずれかから予測時間マップを自動的に知る可能性を容易にする。

【0061】

通常、ハードウェア実装およびネットワーク通信によって、さらなる待ち時間を強いられる。対話型セットアップにおけるこの待ち時間を補正することは、(無線のオーディオ/ビデオストリーミング同期に見られるような)読取りヘッドの単純な置換に変えることがで

10

20

30

40

50

きない。そのような待ち時間の値は、100ミリ秒から、人間の耳の許容可能な心理音響限界をはるかに超える1秒に変動する可能性がある。同期戦略は、この値を任意選択で入力として受け取り、対話型コンテキストに基づいてすべての出力予測を予想する。結果として、(大部分のBluetoothおよびAirMediaストリーミングフォーマットに対応する300msの中域の)比較的小さい値の待ち時間については、ユーザは演奏に先立ってラグを調整する必要はない。本明細書で「物理的時間」の対語として「音楽的時間」で表現する一般的な手法は、そのようなパラメータの自動調整を可能にする。

【0062】

より一般的には、本開示は、実施形態の例として上記で提示した詳細な特徴に限定されず、本開示はさらなる実施形態を包含する。

10

【0063】

一般的には、「ラウドスピーカー」で「伴奏を演奏する」に関係する表現および「あらかじめ記録された楽曲伴奏」の概念は、広く解釈されるものとする。事実上、方法は、たとえばオーディオおよびビデオを含む、どんな「連続的」メディアにも適用される。実際、ビデオ+オーディオコンテンツは、上記で提示した同じ方法を使用して同様に同期することができる。一般的には、上述の「ラウドスピーカー」は、オーディオ-ビデオプロジェクションに置き換えることができ、ビデオフレームはしたがって、同期のために単に予測の位置出力に基づいて上記で提示したように補間され得る。

20

30

40

50

【図面】

【図 1】

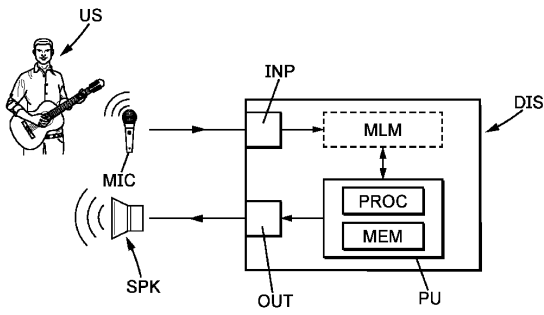
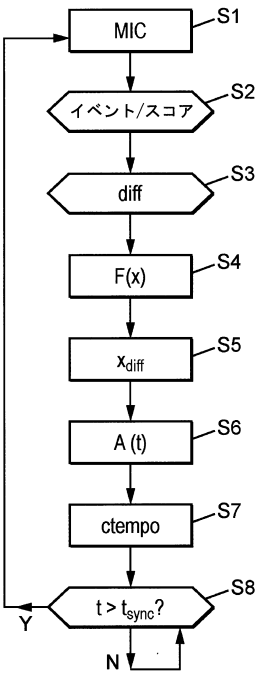
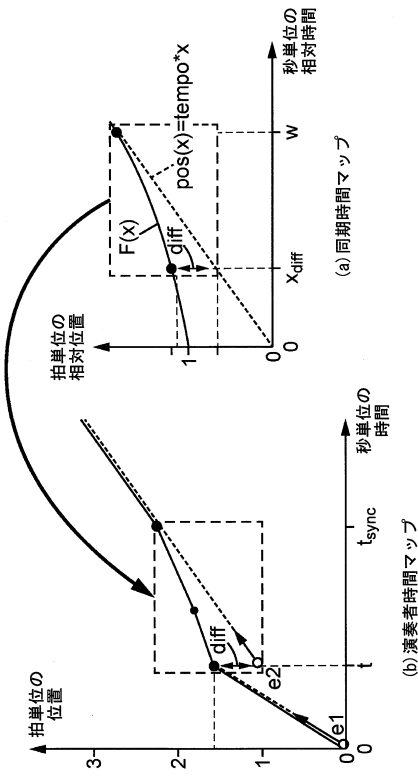


FIG. 1

【図 2】



【図 3】



10

20

30

40

50

フロントページの続き

アール・サン・ミッシェル・ 8 3

(72)発明者 アルシア・コント

フランス・ 9 4 2 5 0 ・ジャンティイ・アヴニユ・パスツール・ 2 6

審査官 上田 雄

(56)参考文献

特開 2 0 0 8 - 1 9 7 2 7 0 (J P , A)

特開 2 0 1 9 - 1 4 4 3 6 0 (J P , A)

特開 2 0 0 1 - 3 1 2 2 7 6 (J P , A)

特開平 0 4 - 1 3 3 0 9 5 (J P , A)

特開 2 0 1 6 - 0 9 9 5 1 2 (J P , A)

特表 2 0 0 2 - 5 1 0 4 0 3 (J P , A)

特開 2 0 0 7 - 2 4 1 1 8 1 (J P , A)

特表平 0 8 - 5 1 0 8 4 6 (J P , A)

国際公開第 2 0 1 8 / 0 1 6 5 8 1 (W O , A 1)

米国特許出願公開第 2 0 1 4 / 0 3 6 0 3 4 0 (U S , A 1)

米国特許第 5 5 2 1 3 2 4 (U S , A)

米国特許第 5 9 1 3 2 5 9 (U S , A)

(58)調査した分野 (Int.Cl. , D B 名)

G 1 0 H 1 / 0 0 - 7 / 1 2

G 1 0 G 1 / 0 0 - 3 / 0 4

G 1 0 L 2 5 / 5 1