

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7610038号
(P7610038)

(45)発行日 令和7年1月7日(2025.1.7)

(24)登録日 令和6年12月23日(2024.12.23)

(51)国際特許分類		F I	
A 6 3 F	13/67 (2014.01)	A 6 3 F	13/67
A 6 3 F	13/30 (2014.01)	A 6 3 F	13/30
A 6 3 F	13/497 (2014.01)	A 6 3 F	13/497
A 6 3 F	13/86 (2014.01)	A 6 3 F	13/86
G 0 6 N	20/00 (2019.01)	G 0 6 N	20/00
請求項の数 19 (全15頁) 最終頁に続く			
(21)出願番号	特願2023-566928(P2023-566928)	(73)特許権者	310021766
(86)(22)出願日	令和4年4月29日(2022.4.29)		株式会社ソニー・インタラクティブエン
(65)公表番号	特表2024-519577(P2024-519577		タテインメント
	A)		東京都港区港南1丁目7番1号
(43)公表日	令和6年5月17日(2024.5.17)	(74)代理人	100105924
(86)国際出願番号	PCT/US2022/026947		弁理士 森下 賢樹
(87)国際公開番号	WO2022/235505	(72)発明者	クリシュナン、ラティシュ
(87)国際公開日	令和4年11月10日(2022.11.10)		アメリカ合衆国、カリフォルニア州 9
審査請求日	令和5年10月30日(2023.10.30)		4 4 0 4、サン マテオ、ブリッジポイ
(31)優先権主張番号	17/307,971		ント パークウェイ 2 2 0 7
(32)優先日	令和3年5月4日(2021.5.4)	(72)発明者	シャー、マウリク
(33)優先権主張国・地域又は機関	米国(US)		アメリカ合衆国、カリフォルニア州 9
			4 4 0 4、サン マテオ、ブリッジポイ
			ント パークウェイ 2 2 0 7
		審査官	鈴木 崇雅
			最終頁に続く

(54)【発明の名称】 機械学習を使用するビデオフレームからのコントローラーアクション認識

(57)【特許請求の範囲】

【請求項1】

一時的信号ではなく、
ビデオフレームのシーケンスを含む記録されたコンピューターシミュレーションを受信することと、
機械学習（ML）モデルに前記ビデオフレームの前記シーケンスを入力することと、
前記MLモデルから出力された、前記記録されたコンピューターシミュレーションを生成することに関連付けられた少なくともいくつかのコンピューターシミュレーションコントローラー（CSC）操作の識別を受信することと、
前記MLモデルから受信された前記CSC操作のうちの少なくとも1つの少なくとも1つのインジケーションとともに、前記記録されたコンピューターシミュレーションを少なくとも1つのオーディオビデオ（AV）ディスプレイに提示することと、
のために、少なくとも1つのプロセッサによって実行可能な命令を含む少なくとも1つのコンピューターメモリ、
を備える、デバイス。

【請求項2】

前記MLモデルは、少なくとも1つのリカレントニューラルネットワーク（RNN）を含む、請求項1に記載のデバイス。

【請求項3】

前記RNNは、少なくとも1つの長短期記憶（LSTM）ネットワークを含む、請求項

2 に記載のデバイス。

【請求項 4】

前記 M L モデルは、少なくとも 1 つの畳み込みニューラルネットワーク（C N N）を含む、請求項 1 に記載のデバイス。

【請求項 5】

前記少なくとも 1 つのプロセッサを含み、

前記少なくとも 1 つのプロセッサは前記 A V ディスプレイに組み込まれている、請求項 1 に記載のデバイス。

【請求項 6】

前記少なくとも 1 つのプロセッサを含み、

前記少なくとも 1 つのプロセッサは前記コンピューターシミュレーションのソースに組み込まれている、請求項 1 に記載のデバイス。

【請求項 7】

前記ソースは、少なくとも 1 つのコンピューターシミュレーションコンソールを含む、請求項 6 に記載のデバイス。

【請求項 8】

前記ソースは、ワイドエリアコンピューターネットワークを通じて前記 A V ディスプレイと通信する少なくとも 1 つのサーバーを含む、請求項 6 に記載のデバイス。

【請求項 9】

少なくとも 1 つのコンピューターシミュレーションコントローラーの制御下で生成された、少なくとも 1 つの記録されたコンピューターシミュレーションの少なくともビデオを提示するように構成された少なくとも 1 つのディスプレイであって、前記少なくとも 1 つの記録されたコンピューターシミュレーションは前記少なくとも 1 つの記録されたコンピューターシミュレーションの前記ビデオの生成中の前記コンピューターシミュレーションコントローラーの操作に関する情報を含まない、少なくとも 1 つのディスプレイと、
少なくとも 1 つのプロセッサと、を備え、

前記少なくとも 1 つのプロセッサが、

前記少なくとも 1 つの記録されたコンピューターシミュレーションの前記ビデオの生成中に、少なくとも 1 つの機械学習（M L）モデルに前記少なくとも 1 つの記録されたコンピューターシミュレーションを入力することにより前記 M L モデルから出力された、前記

コンピューターシミュレーションコントローラーの操作に関する情報を受信することと、
前記少なくとも 1 つの記録されたコンピューターシミュレーションの前記ビデオを提示することとともに前記情報を提示するために、前記少なくとも 1 つの記録されたコンピューターシミュレーションの前記ビデオの生成中の前記コンピューターシミュレーションコントローラーの操作に関する前記情報を、前記少なくとも 1 つのディスプレイに提示することと、

を実行する、装置。

【請求項 10】

前記プロセッサは、前記ディスプレイに組み込まれている、請求項 9 に記載の装置。

【請求項 11】

前記プロセッサは、コンピューターシミュレーションコンソールに組み込まれている、請求項 9 に記載の装置。

【請求項 12】

前記プロセッサは、ワイドエリアネットワークを通じて前記ディスプレイと通信するサーバーに組み込まれている、請求項 9 に記載の装置。

【請求項 13】

前記 M L モデルは、少なくとも 1 つのリカレントニューラルネットワーク（R N N）を含む、請求項 9 に記載の装置。

【請求項 14】

前記 R N N は、少なくとも 1 つの長短期記憶（L S T M）ネットワークを含む、請求項

10

20

30

40

50

13に記載の装置。

【請求項15】

少なくとも1つの機械学習（ML）モデルに、少なくともトレーニングセットを入力することであって、前記トレーニングセットは、複数の記録されたコンピューターシミュレーションからのビデオフレームのシーケンスと、前記ビデオフレームのシーケンスの生成中に実行されたコンピューターシミュレーションコントローラ（CSC）操作に関する前記ビデオフレームのシーケンスに関連付けられた情報とを含む、前記入力することと、

前記MLモデルに、第1の記録されたコンピューターシミュレーションの生成中に実行されたCSC操作に関する情報を含まない少なくとも前記第1の記録されたコンピューターシミュレーションを入力することと、

前記第1の記録されたコンピューターシミュレーションの生成中に実行されたCSC操作に関する情報であって、前記MLモデルから出力された前記CSC操作に関する情報を受信することと、

前記CSC操作に関する情報とともに、前記第1の記録されたコンピューターシミュレーションを提示することと、

を含む、方法。

【請求項16】

前記MLモデルから受信された前記第1の記録されたコンピューターシミュレーションの生成中に実行された前記CSC操作に関する情報を、前記第1の記録されたコンピューターシミュレーションを視覚的に提示することとともに、聴覚的に提示することを含む、請求項15に記載の方法。

【請求項17】

前記MLモデルから受信された前記第1の記録されたコンピューターシミュレーションの生成中に実行された前記CSC操作に関する情報を、前記第1の記録されたコンピューターシミュレーションを視覚的に提示することとともに、視覚的に提示することを含む、請求項15に記載の方法。

【請求項18】

前記第1の記録されたコンピューターシミュレーションを提示するディスプレイに前記第1の記録されたコンピューターシミュレーションを提供するサーバーにおいて、前記第1の記録されたコンピューターシミュレーションを、前記第1の記録されたコンピューターシミュレーションの生成中に実行されたCSC操作に関する情報と関連付けることを含む、請求項15に記載の方法。

【請求項19】

前記第1の記録されたコンピューターシミュレーションを提示するディスプレイに前記第1の記録されたコンピューターシミュレーションを提供するローカルソースにおいて、前記第1の記録されたコンピューターシミュレーションを、前記第1の記録されたコンピューターシミュレーションの生成中に実行されたCSC操作に関する情報と関連付けることを含む、請求項15に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本願は、必然的にコンピューター技術が原因となり、具体的な技術的改善を生じさせる、技術的に発明性のある非定型な解決策に関する。

【背景技術】

【0002】

本明細書で理解されるように、以前にプレイしたコンピューターゲームのビデオは、コンピューターネットワークを通じて共有され得、ゲームのレベルを完了することによる等、視聴者にゲームに成功する方法を指導する。さらに本明細書で理解されるように、そのようなゲームビデオは、どのコントローラボタンが押されたのかに関する情報及びゲームのプレイ中のタイミングに関する情報を含まない場合がある。その理由として、ゲーム

10

20

30

40

50

ビデオは、ゲームがプレイされたときにコントローラーアクションをキャプチャすることなく記録され得るためである。

【発明の概要】

【0003】

また、本明細書で理解されるように、どのコントローラーボタンが押されたのかに関する情報、及びコンピューターゲームのプレイを学習するプレイヤーに価値があるタイミングに関する情報は、初心者からスピードランナーまで、多くのタイプのゲーマーにとって、ゲームをより楽しいものにする。本明細書では、追加のコントローラーデータを用いずに一連のビデオフレームを分析することによって、コントローラーアクション情報を生成する機械学習技術が提供される。

10

【0004】

したがって、デバイスは少なくとも1つのコンピューターメモリを含み、少なくとも1つのコンピューターメモリは、一時的信号ではなく、さらには、ビデオフレームのシーケンスを含む記録されたコンピューターシミュレーションを受信するために、少なくとも1つのプロセッサによって実行可能な命令を含む。命令は、機械学習（ML）モデルにおいてビデオフレームのシーケンスを処理すること、記録されたコンピューターシミュレーションを生成することに関連付けられたコンピューターシミュレーションコントローラー（CSC）操作の識別をMLモデルから受信することを実行可能にする。加えて、命令は、記録されたコンピューターシミュレーションを少なくとも1つのオーディオビデオ（AV）ディスプレイに、MLモデルから受信されたCSC操作の少なくとも1つのインジケーションとともに提示することを実行可能にする。

20

【0005】

例示的な実施形態では、MLモデルは、少なくとも1つの長短期記憶（LSTM）ネットワーク等の少なくとも1つのリカレントニューラルネットワーク（RNN）を含む。また、畳み込みニューラルネットワーク（CNN）も使用できる。

【0006】

本デバイスはプロセッサを含み得、このプロセッサは、AVディスプレイに、もしくはコンピューターシミュレーションコンソール等のコンピューターシミュレーションのソースに、及び／またはワイドエリアコンピューターネットワークを通じてAVディスプレイと通信するサーバーに、組み込まれ得る。

30

【0007】

別の態様では、本装置は、少なくとも1つのコンピューターシミュレーションコントローラーの制御下で生成された、少なくとも1つの記録されたコンピューターシミュレーションのビデオを提示するように構成された少なくとも1つのディスプレイを含む。しかしながら、記録されたコンピューターシミュレーションは、少なくとも1つの記録されたコンピューターシミュレーションのビデオの生成中におけるコンピューターシミュレーションコントローラーの操作に関する情報を含まない。したがって、本装置は少なくとも1つのプロセッサを含み、プロセッサは、少なくとも1つの記録されたコンピューターシミュレーションのビデオの生成中に、ビデオから、コンピューターシミュレーションコントローラーの操作に関する情報を識別するための命令を伴って構成される。命令は、少なくとも1つの記録されたコンピューターシミュレーションのビデオを提示することとともに当該情報を提示するために、少なくとも1つの記録されたコンピューターシミュレーションのビデオの生成中のコンピューターシミュレーションコントローラーの操作に関する情報を、少なくとも1つのディスプレイに提供するために実行可能である。

40

【0008】

例示的な実施態様では、少なくとも1つの機械学習（ML）モデルを使用して、少なくとも1つの記録されたコンピューターシミュレーションのビデオの生成中に、ビデオから、コンピューターシミュレーションコントローラーの操作に関する情報を識別するために実行可能であり得る。

【0009】

50

別の態様では、本方法は、少なくともトレーニングセットを、少なくとも1つの機械学習（ML）モデルに入力することを含む。トレーニングセットは、複数の記録されたコンピュータシミュレーションからのビデオフレームのシーケンスと、ビデオフレームのシーケンスの生成中に実行されたコンピュータシミュレーションコントローラ（CSC）操作に関するビデオフレームのシーケンスに関連付けられた情報とを含む。本方法は、次に、少なくとも第1の記録されたコンピュータシミュレーションをMLモデルに入力することを含み、第1の記録されたコンピュータシミュレーションは、第1の記録されたコンピュータシミュレーションの生成中に実行されたCSC操作に関する情報を含まない。本方法は、MLモデルから受信された第1の記録されたコンピュータシミュレーションの生成中に実行されたCSC操作に関する可聴情報及び／または可視情報とともに第1の記録されたコンピュータシミュレーションを提示することを含む。

10

【0010】

本願の詳細は、その構造及び動作の両方について、添付の図面を参照すると最良に理解でき、図面において、類似の参照符号は、類似の部分を指す。

【図面の簡単な説明】

【0011】

【図1】本原理による例を含む例示的なシステムのブロック図である。

【図2】本原理に即した記録されたコンピュータシミュレーション（コンピュータゲーム等）のビデオから学習され得る例示的なキー操作を要するゲームコントローラの例を示す。

20

【図3】本原理に即したコンピュータゲーム等の記録されたコンピュータシミュレーションを閲覧するための例を示す。

【図4】どのようなコントローラ操作が実施され、いつ実施されたかに関する情報を、記録されたコンピュータシミュレーションビデオから導出するために、機械学習（ML）モデルをトレーニングするための例示的なフローチャートフォーマットにおける例示的なロジックを示す。

【図5】MLモデルをトレーニングするために使用され得るコントローラ操作に関する付随情報を伴う一連のビデオフレームを概略的に示す。

【図6】コントローラ操作に依存して発生する可能性があるコントローラ操作及び代替ビデオパスに関する付随情報を伴わないビデオフレームのシーケンスを概略的に示し、コントローラ操作の識別を示す。

30

【図7】記録されたコンピュータシミュレーションからの記録されたビデオフレームのシーケンスからのコントローラ操作を識別するための、例示的なフローチャートフォーマットの例示的なロジックを示す。

【図8】本原理に即したMLモデルからの例示的な出力表現を示す。

【発明を実施するための形態】

【0012】

したがって、本原理では、深層学習モデルを含む機械学習モデルを使用し得る。機械学習モデルは、教師あり学習、教師なし学習、半教師あり学習、強化学習、機能学習、自己学習、及び他の学習形式を含む方法でトレーニングされた様々なアルゴリズムを使用する。コンピュータ回路によって実装できるそのようなアルゴリズムの例は、畳み込みニューラルネットワーク（CNN）、一連の画像から情報を学習するのに適し得るリカレントニューラルネットワーク（RNN）、及び長短期記憶（LSTM）ネットワークとして既知のRNNのタイプ等の1つ以上のニューラルネットワークを含む。サポートベクターマシン（SVM）及びベイジアンネットワークは、また、機械学習モデルの例と見なされ得る。

40

【0013】

本明細書で理解されるように、機械学習を行うことは、トレーニングデータでモデルにアクセスして、次にトレーニングし、モデルがさらなるデータを処理して予測を行うのを可能にすることを含む。ニューラルネットワークは、入力層、出力層、及びそれらの間に

50

複数の隠れ層を含み得、それらの層は、適切な出力について推論するように構成及び重み付けされる。

【0014】

本開示は、概して、限定ではないが、コンピューターゲームネットワーク等の家電製品（CE）デバイスネットワークの態様を含むコンピューターエコシステムに関する。本明細書のシステムはネットワークを通じて接続され得るサーバーコンポーネント及びクライアントコンポーネントを含み得、それにより、クライアントコンポーネントとサーバーコンポーネントとの間でデータが交換され得る。クライアントコンポーネントは、Sony PlayStation（登録商標）等のゲーム機またはMicrosoft（登録商標）もしくはNintendo（登録商標）もしくは他の製造者によって作成されたゲーム機、仮想現実（VR）ヘッドセット、拡張現実（AR）ヘッドセット、ポータブルテレビ（例えば、スマートテレビ、インターネット対応テレビ）、ラップトップ及びタブレットコンピューター等のポータブルコンピューター、ならびにスマートフォン及び下記に説明される追加例を含む他のモバイルデバイスを含む、1つ以上のコンピューティングデバイスを含み得る。これらのクライアントデバイスは、様々な動作環境で動作し得る。例えば、クライアントコンピューターの一部は、例として、Linux（登録商標）オペレーティングシステム、Microsoft（登録商標）のオペレーティングシステム、もしくはUnix（登録商標）オペレーティングシステム、またはApple（登録商標）社もしくはGoogle（登録商標）によって製造されたオペレーティングシステムを使用し得る。これらの動作環境を使用して、Microsoft（登録商標）もしくはGoogle（登録商標）もしくはMozilla（登録商標）によって作成されたブラウザ、または下記に説明されるインターネットサーバーによってホストされるウェブサイトにアクセスできる他のブラウザプログラム等の1つ以上の閲覧プログラムを実行し得る。また、本原理に従った動作環境を使用して、1つ以上のコンピューターゲームプログラムを実行し得る。

【0015】

サーバー及び／またはゲートウェイは、インターネット等のネットワークを通じてデータを受信及び伝送するサーバーを構成する命令を実行する1つ以上のプロセッサを含み得る。または、クライアント及びサーバーは、ローカルイントラネットまたは仮想プライベートネットワークを通じて接続できる。サーバーまたはコントローラーは、Sony PlayStation（登録商標）等のゲーム機、パーソナルコンピューター等によってインスタンス化され得る。

【0016】

情報は、クライアントとサーバーとの間でネットワークを通じて交換され得る。この目的のために及びセキュリティのために、サーバー及び／またはクライアントは、ファイアウォール、ロードバランサ、テンポラリストレージ、及びプロキシ、ならびに信頼性及びセキュリティのための他のネットワークインフラストラクチャを含み得る。1つ以上のサーバーは、ネットワークメンバーにオンラインソーシャルウェブサイト等のセキュアコミュニティを提供する方法を実施する装置を形成し得る。

【0017】

プロセッサは、アドレスライン、データライン、及び制御ライン等の様々なライン、ならびにレジスタ及びシフトレジスタによってロジックを実行できるシングルチッププロセッサまたはマルチチッププロセッサであり得る。

【0018】

一実施形態に含まれるコンポーネントは、他の実施形態では、任意の適切な組み合わせで使用できる。例えば、本明細書に説明される及び／または図で示される様々なコンポーネントのいずれかは、組み合わされ得る、交換され得る、または他の実施形態から排除され得る。

【0019】

「A、B、及びCのうちの少なくとも1つを有するシステム」（同様に「A、B、また

10

20

30

40

50

はCのうちの少なくとも1つを有するシステム」及び「A、B、Cのうちの少なくとも1つを有するシステム」)は、A単独、B単独、C単独、A及びBをともに、A及びCをともに、B及びCをともに、ならびに／またはA、B、及びC等をともに有するシステムを含む。

【0020】

ここで具体的に図1を参照すると、例示的なシステム10が示され、システム10は、本原理による、上述され及び下記に詳述される、例示的なデバイスの1つ以上を含み得る。システム10に含まれる第1の例示的デバイスは、限定ではないが、テレビチューナ(同様に、テレビを制御するセットトップボックス)を有するインターネット対応テレビ等の音声ビデオデバイス(AVD)12等の家電製品(CE)デバイスである。代わりに、AVD12は、また、コンピューター制御型インターネット対応(「スマート」)電話、タブレットコンピューター、ノートブックコンピューター、HMD、ウェアラブルコンピューター制御デバイス、コンピューター制御型インターネット対応ミュージックプレイヤー、コンピューター制御型インターネット対応ヘッドフォン、インプラント可能な皮膚用デバイス等のコンピューター制御型インターネット対応のインプラント可能なデバイス等であり得る。とにかく、AVD12は、本原理を実施するように構成される(例えば、本原理を実施するように他のCEデバイスと通信し、本明細書に説明されるロジックを実行し、本明細書に説明されるいずれかの他の機能及び／または動作を行う)ことを理解されたい。

【0021】

したがって、そのような原理を実施するために、AVD12は、図1に示されるコンポーネントの一部または全てによって確立できる。例えば、AVD12は、1つ以上のディスプレイ14を含み得、1つ以上のディスプレイ14は、高解像度もしくは超解像度の「4K」またはより高い解像度のフラットスクリーンによって実装され得、ディスプレイのタッチによりユーザー入力信号を受信するためにタッチ対応であり得る。AVD12は、本原理に従って音声を出力するための1つ以上のスピーカー16と、可聴コマンドをAVD12に入力して、AVD12を制御するための、例えば、音声受信機／マイクロホン等の少なくとも1つの追加入力デバイス18と、を含み得る。例示的なAVD12は、また、1つ以上のプロセッサ24の制御の下、インターネット、WAN、LAN等の少なくとも1つのネットワーク22を通じて通信するための1つ以上のネットワークインターフェース20を含み得る。また、グラフィックプロセッサを含み得る。したがって、インターフェース20は、限定ではないが、Wi-Fi(登録商標)送受信機であり得、Wi-Fi(登録商標)送受信機は、限定ではないが、メッシュネットワークトランシーバー等の無線コンピューターネットワークインターフェースの例である。プロセッサ24は、ディスプレイ14をそこで画像を提示するように制御することと、そこから入力を受信すること等の本明細書に説明されるAVD12の他の要素を含む本原理を実施するようにAVD12を制御することを理解されたい。さらに、ネットワークインターフェース20は、有線もしくは無線のモデムもしくはルータ、または、無線テレフォニ送受信機もしくは上述したWi-Fi(登録商標)送受信機等の他の適切なインターフェースであり得ることに留意されたい。

【0022】

上述に加えて、AVD12は、また、別のCEデバイスに物理的に接続するための高精細マルチメディアインターフェース(HDMI(登録商標))ポートもしくはUSBポート、及び／またはヘッドフォンを経由してAVD12からユーザーに音声を提示するためにAVD12にヘッドフォンを接続するためのヘッドフォンポート等の1つ以上の入力ポート26を含み得る。例えば、入力ポート26は、オーディオビデオコンテンツのケーブルまたは衛星ソース26aに有線または無線で接続され得る。したがって、ソース26aは、分離もしくは統合されたセットトップボックス、またはサテライト受信機であり得る。または、ソース26aは、コンテンツを含むゲーム機またはディスクプレイヤーであり得る。ソース26aは、ゲーム機として実装されるとき、CEデバイス44に関連して

下記に説明されるコンポーネントの一部または全てを含み得る。

【0023】

AVD12は、さらに、一時的信号ではない、ディスクベースストレージまたはソリッドステートストレージ等の1つ以上のコンピューターメモリ28を含み得、これらのストレージは、いくつかの場合、スタンドアロンデバイスとして、またはAVプログラムを再生するためにAVDのシャーシの内部もしくは外部のいずれかでパーソナルビデオ録画デバイス(PVR)もしくはビデオディスクプレイヤーとして、または取り外し可能メモリ媒体として、AVDのシャーシ内で具現化される。また、いくつかの実施形態では、AVD12は、限定ではないが、サテライト基地局もしくは携帯電話基地局から地理的位置情報を受信し、情報をプロセッサ24に提供し、及び／またはAVD12がプロセッサ24と併せて配置される高度を判定するように構成される、携帯電話受信機、GPS受信機、及び／または高度計30等の位置受信機または場所受信機を含み得る。コンポーネント30は、また、3次元でのAVD12の位置及び向きを判定するために、典型的に、加速度計、ジャイロスコープ、及び磁力計の組み合わせを含む慣性測定ユニット(IMU)によって実現され得る。

10

【0024】

AVD12の説明を続けると、いくつかの実施形態では、AVD12は、1つ以上のカメラ32を含み得、1つ以上のカメラ32は、熱画像カメラ、ウェブカメラ等のデジタルカメラ、ならびに／または本原理に従って写真／画像及び／もしくはビデオを集めるように、AVD12に統合され、プロセッサ24によって制御可能であるカメラであり得る。また、AVD12に含まれるのは、Bluetooth(登録商標)及び／または近距離無線通信(NFC)技術を各々使用して、他のデバイスと通信するためのBluetooth(登録商標)送受信機34及び他のNFC要素36であり得る。例示的なNFC素子は、無線自動識別(RFID)素子であり得る。

20

【0025】

さらにまた、AVD12は、プロセッサ24に入力を提供する1つ以上の補助センサ38(例えば、加速度計、ジャイロスコープ、サイクロメータ等の運動センサ、または磁気センサ、赤外線(IR)センサ、光学センサ、速度センサ及び／またはケイデンスセンサ、ジェスチャセンサ(例えば、ジェスチャコマンドを検知するためのセンサ))を含み得る。AVD12は、プロセッサ24に入力を提供する無線(over-the-air(OTA))TV放送を受信するための無線テレビ放送ポート40を含み得る。上述に加えて、AVD12は、また、赤外線(IR)データアソシエーション(IRDA)デバイス等のIR伝送機及び／またはIR受信機及び／またはIR送受信機42も含み得ることに留意されたい。バッテリー(図示せず)は、AVD12に給電するために提供され得、運動エネルギーを電力に変換してバッテリーを充電し、及び／またはAVD12に給電し得る運動エネルギーハーベスターであり得る。グラフィックスプロセッシングユニット(GPU)44及びフィールドプログラマブルゲートアレイ46も含み得る。

30

【0026】

さらに図1を参照すると、AVD12に加えて、システム10は、1つ以上の他のCEデバイスタイプを含み得る。一例では、第1のCEデバイス48は、AVD12に直接送信されたコマンドを介して及び／または後述のサーバーを経由して、コンピューターゲームの音声及びビデオをAVD12に送信するために使用できるコンピューターゲーム機であり得る一方、第2のCEデバイス50は第1のCEデバイス48と同様のコンポーネントを含み得る。示される例では、第2のCEデバイス50は、プレイヤーによってマニピュレートされたコンピューターゲームコントローラー、またはプレイヤーによって装着されたヘッドマウントディスプレイ(HMD)として構成され得る。示される例では、2つのCEデバイスだけが示され、より少ない数またはより多い数のデバイスが使用され得ることを理解されたい。本明細書のデバイスは、AVD12用に示されるコンポーネントの一部または全てを実装し得る。次の図に示されるコンポーネントのいずれかは、AVD12の場合に示されるコンポーネントの一部または全てを組み込み得る。

40

50

【0027】

ここで、上述の少なくとも1つのサーバー52を参照すると、サーバー52は、少なくとも1つのサーバープロセッサ54と、ディスクベースストレージまたはソリッドステートストレージ等の少なくとも1つの有形コンピューター可読記憶媒体56と、サーバープロセッサ54の制御の下、ネットワーク22を通じて図1の他のデバイスとの通信を可能にし、実際に、本原理に従ってサーバーとクライアントデバイスとの間の通信を容易にし得る少なくとも1つのネットワークインターフェース58とを含む。ネットワークインターフェース58は、例えば、有線もしくは無線のモデムもしくはルータ、Wi-Fi（登録商標）送受信機、または、例えば、無線テレフォニ送受信機等の他の適切なインターフェースであり得ることに留意されたい。

10

【0028】

したがって、いくつかの実施形態では、サーバー52は、インターネットサーバーまたはサーバー「ファーム」全体であり得、「クラウド」機能を含み得、「クラウド」機能を行い得、それにより、システム10のデバイスは、例えば、ネットワークゲーミングアプリケーションに関する例示的な実施形態ではサーバー52を介して「クラウド」環境にアクセスし得る。または、サーバー52は、1つ以上のゲーム機、または図1に示される他のデバイスと同じ部屋もしくはその近くにある他のコンピューターによって実装され得る。

【0029】

以下の図に示されるコンポーネントは、図1に示されるコンポーネントの一部または全てを含み得る。

20

【0030】

図2は、ゲームのプレイ中にコンピューターゲーム等のコンピューターシミュレーションを制御するために本原理に従って使用され得るゲームコントローラーの非限定的な例を示し、ゲームに関連付けられたゲームのビデオ（及びオーディオ／触覚等）は、ディスプレイ14等のディスプレイに示され（触覚の場合、例えばコントローラーを使用して触覚検出のために生成され）、また再生のために記録され、例えばコンピューターメモリ28は図1に示される、及び／またはサーバーメモリ56は図1に示される。

【0031】

ゲームコントローラーは、ディスプレイ14に提示されたコンピューターゲームを制御するために、コンピューターシミュレーションのソースと通信する上述したコンポーネント（コンピューターゲーム機として組み込まれたCEデバイス48及び／またはサーバー52等）のうちの1つ以上を組み込むことができると理解されたい。

30

【0032】

図2は、円形のほぼ円筒形状の左右のハンドル202，204を伴う軽量のハンドヘルド筐体を含むコントローラー200を示し、各ハンドルは上面を規定しており、上面には4つの操作可能なキーが配置されている。例えば4つの方向キー206は、左側のハンドル202の上部に十字形パターンに並んでいる。キー206を使用して、ディスプレイの各々の方向にオブジェクトを移動させることができる。

【0033】

左ハンドル202のすぐ前方に、追加のL1及びL2キー208が提供され得る。ブリッジはハンドル202，204を接続し、セレクトキー210は、スタートキー212とともにブリッジ上に配置され得る。

40

【0034】

右ハンドル204の4つのキーは、三角形キー214、正方形キー216、及び「O」キー218、及び「X」キー220を含み得、各キーは、ゲーム設計者の要望に従って各々の機能を担い得る。右ハンドル204のすぐ前方に、追加のR1及びR2キー222が提供され得る。

【0035】

また、ハンドル202とハンドル204との間に、左ハンドル202のすぐ内側に左ジョイスティック224が提供され得る。左ジョイスティック224は押し下げ可能な上部

50

226を含み得る。同様に、右ジョイスティック228は右ハンドル204のすぐ内側に提供され得る。右ジョイスティック228は押し下げ可能な上部230を含み得る。

【0036】

図3はデバイス300を示し、デバイス300は、コンピューターシミュレーション（コンピューターゲーム等）のビデオを記録し、必要に応じて、そのビデオに関連付けられたオーディオ及び触覚を記録する。デバイス300は、例えば、ビデオを生成するためにゲームがプレイされるデバイス、コンピューターゲーム機、プレイヤーから離れたコンピューターゲームサーバー、またはそれらの組み合わせによって実装され得る。

【0037】

記録されたコンピューターシミュレーション（コンピューターゲーム等）は、記録されたコンピューターゲームのソース302に提供される。ソース302は、記録されたビデオ（必要に応じて、記録されたオーディオ及びゲームの他の感覚的出力、本明細書ではオーディオビデオと呼ぶ）を、観客／学習者コンピューターデバイス304に提供する。

【0038】

記録中にゲームを制御するために入力されたコンピューターゲームコントローラーの操作を記録することなく、コンピューターゲームAVを記録し得るため、機械学習（ML）エンジン306は、記録されたゲームAVとともに観客／コンピューター304にインジェクションを提示するために、コントローラー操作をコンピューターゲームオーディオビデオから再現するために、本明細書の原理に従った実行のために提供される。MLモデル306は、少なくとも1つの長短期記憶（LSTM）ネットワーク等の少なくとも1つのリカレントニューラルネットワーク（RNN）を含み得る。また、畳み込みニューラルネットワーク（CNN）も使用し得る。MLモデル306は、観客／学習者コンピューター304、記録されたゲームソース302（リモートサーバーまたはローカルゲーム機等）におけるプロセッサ等、本明細書で開示されるプロセッサのいずれかまたはそれらの組み合わせによって実行され得る。いくつかの実施形態では、要素300、302、及び304は同一のデバイスによって実装され得る。例えば、ユーザーは、同じコンソール上でより早くプレイしていた異なるユーザーのゲーム記録から学習しようとし得る。

【0039】

図4は、図3のMLモデル306をトレーニングするためのロジックを示す。ブロック400において、データのトレーニングセットはMLモデルに入力される。訓練セットは、コンピューターゲーム等の典型的に複数のコンピューターシミュレーションに関連付けられたビデオ及び／またはオーディオを含み得る。

【0040】

より高い特異性により、トレーニングセットは、複数の記録されたコンピューターシミュレーションからのビデオフレームのシーケンス、及び／または必要に応じて、ビデオに関連付けられたオーディオを含み得る。さらに、トレーニングセットは、ビデオフレームのシーケンスの発生中に実行されたコンピューターシミュレーションコントローラー（CSC）操作に関するビデオフレームのシーケンスに関連付けられたグラウンドトゥールズ情報を含む。CSC操作は、例えば、図2の例に示される制御のうちの任意の1つ以上のマニピュレーションから生じ得る。トレーニングセットから、及び適切な学習技術（例えば、教師あり、教師なし、半教師あり等）を使用して、ブロック402において、MLモデル306はビデオフレームシーケンスからCSC操作を学習する。

【0041】

図5は、MLモデルをトレーニングするために使用されるトレーニングセットにおけるビデオフレーム502のシーケンス500を示す。メタデータ504は、必要に応じて、フレームごとにシーケンス500に関連付けられ、トレーニングシーケンスの記録中にどのコントローラー操作が発生したかと、そのコントローラー操作がいつ発生したかを示す。いくつかの実施形態では、メタデータはフレーム毎である必要はないことに留意されたい。メタデータは、CSC操作が以前のフレームと比較して変化していた場合にのみ更新され得る。また、CSC操作は、ビデオデータ及びCSCデータを同期させるために使

10

20

30

40

50

用できたタイムスタンプとともに、完全に別個のファイルに記憶できる。メタデータ504は、例えば図2に示されるコントローラー等のコントローラーにおいて、どのような特定の制御面がマニピュレートされたのかのインジケーションと、これらの制御面が記録されたシーケンス500の生成中にいつマニピュレートされたかのインジケーションとを含み得る。

【0042】

トレーニングセットは様々な方法で作成され得る。ボタンマニピュレーションのグラウンドトゥールズはゲームプレイ中に集められ、トレーニングセットの一部として使用するために、時間整合して生成されたビデオに関連付けられ得る。トレーニングセットは、また、予め記録されたゲームビデオも含み得、これらのゲームビデオは、それらの上にコントローラーオーバーレイを有し、コントローラーオーバーレイは、オリジナルのゲームがプレイされ、ビデオとともに提示されたときに生成される。オーバーレイにおけるコントローラー操作データは、典型的にビデオ上に提示されるため、既にビデオと時間整合している。各フレームのビデオの画素値をチェックして、ラベリングデータを生成するためにオーバーレイで示されたように、どのボタンが押されたかを確認し得る。これにより、各フレームがタイムスタンプに関連付けられ得るので、タイミングデータも与えられる。オーバーレイ機能は、トレーニングセットビデオが生成されるときに初期再生時にオンにできる。

【0043】

グラウンドトゥールズのコントローラー操作は、コントローラーの制御下でゲームビデオを記録するデバイスにコントローラーからストリーミングされ得、これによって、トレーニングセットの要素を確立するために、グラウンドトゥールズのコントローラー操作をビデオに関連付ける。

【0044】

トレーニングに続いて、MLモデルは、第1の記録されたコンピューターシミュレーションの生成中に実行されたCSC操作に関する情報を含まないビデオフレームのシーケンス600（図6では、分かり易くするために、オプションのオーディオを示していない）を伴う第1の記録されたコンピューターシミュレーションを受信するために使用され得る。ビデオフレームのシーケンスがどのように推移するかに応じて、MLモデルは図5のトレーニングに基づき、第1の記録されたシミュレーションの生成中にどのCSC操作が行われたかと、CSC操作がいつ行われたかとを識別する。例えば、図6に示されるシーケンス1と、潜在的なシーケンス2Aとの間で、CSCコントローラー操作が発生していない場合がある。一方、シーケンス1と潜在的なシーケンス2B（潜在的なシーケンス2Aとは異なるシーケンス）との間では、シーケンス全体をシーケンス1 - シーケンス2Aからシーケンス1 - シーケンス2BにずらすようにCSC操作を行い得、そのようなCSC操作の識別602はMLモデルによって行われる。

【0045】

図7は、例示的なフローチャートフォーマットで上述した原理を示す。図7のロジックは、観客／学習者AVディスプレイ304、及び／または生成されたコンピューターシミュレーションのソース300、及び／または記録されたコンピューターシミュレーションのソース302に組み込まれたプロセッサのうちの任意の1つ以上によって実行され得、それらのいずれか一方または両方は、コンピューターシミュレーションコンソール、またはインターネット等の広域コンピューターネットワークを通じてAVディスプレイと通信するサーバーとしてインスタンス化され得る。

【0046】

下記に加えて、MLモデルがCSC操作を識別するのに必要な時間を短くするために、前処理及びダウンスケーリングに対してブロックが提供され得、同様に、サポートされていないCSC操作をフィルタリングし、レイテンシを調整するために後処理及び同期に対してブロックが提供され得ることを理解されたい。

【0047】

ステップ700から開始して、ビデオフレームのシーケンス及び／または付随するオーディオサウンドトラックを含む、記録されたコンピューターゲーム等の記録されたコンピューターシミュレーションを受信する。典型的に、記録されたシミュレーションは、記録されたコンピューターシミュレーションの生成中のコンピューターシミュレーションコントローラーの操作に関する情報を含まない。フレームのシーケンスは、例えば、インターネットプラットフォームからのゲームビデオのスニペットであり得る。

【0048】

ブロック702に進むと、記録されたコンピューターシミュレーション、例えば、記録されたシミュレーションにおけるビデオフレームのシーケンスは、図4を参照して開示されるようにトレーニングされた、図3におけるMLモデル306によって処理される。MLモデルはコンピューターシミュレーションコントローラー(CSC)操作のインジケーションを識別し、出力し、そして、CSC操作は、受信され、記録されたコンピューターシミュレーションに関連付けられたトレーニングに従って記録されたコンピューターシミュレーションの生成に関連付けられる。MLモデルによって識別されたCSC操作は、例えば、図2の例では、どの制御要素がマニピュレートされたかと、いつマニピュレートされたかを含み得る。

【0049】

ブロック704に進むと、記録されたコンピューターシミュレーションは、MLモデルから受信されたCSC操作の少なくとも1つのインジケーションとともに、図3の観客／学習者コンピューター304等の少なくとも1つのオーディオビデオ(AV)ディスプレイに提示される。

【0050】

図8には以下のことが示される。記録されたビデオゲーム800は、図1に示されるディスプレイ14等のディスプレイに提示され得る。図中、ゲーム800は、飛行物体806において武器804を射撃するキャラクタ802を含む。ライン808は、オブジェクト806がヒットしたことを示す。

【0051】

どのCSC操作が発生したか(「赤キーが押された」と、CSC操作がゲーム中にいつ発生したかを示すインジケーション810がディスプレイに提示され、例では「今」で示され、爆発に繋がった過去のCSC操作、及び操作時間を示し、また探索のために今後のCSC操作も示し得ることを理解されたい。インジケーション810は、示されるように視覚的に提示され得る、及び／またはディスプレイに関連付けられたスピーカーで聴覚的に提示され得る。このように、記録されたコンピューターシミュレーションは、記録されたコンピューターシミュレーションの生成中に実行されたCSC操作に関するMLモデルから受信された情報とともに提示される。

【0052】

上述のロジックは、コンピューターゲームコントローラーまたは他の機能を伴うプラグインとして提供されることによって、ゲーマーがゲームビデオをダウンロードし、ビデオを生成したコントローラー操作のシーケンスに関する情報を取得することを可能にし得る。

【0053】

いくつかの例示的な実施形態を参照して本原理を説明してきたが、これらは限定することを意図していないことと、様々な代替の配置を使用して、本明細書で特許請求される主題を実施し得ることとが認識される。

10

20

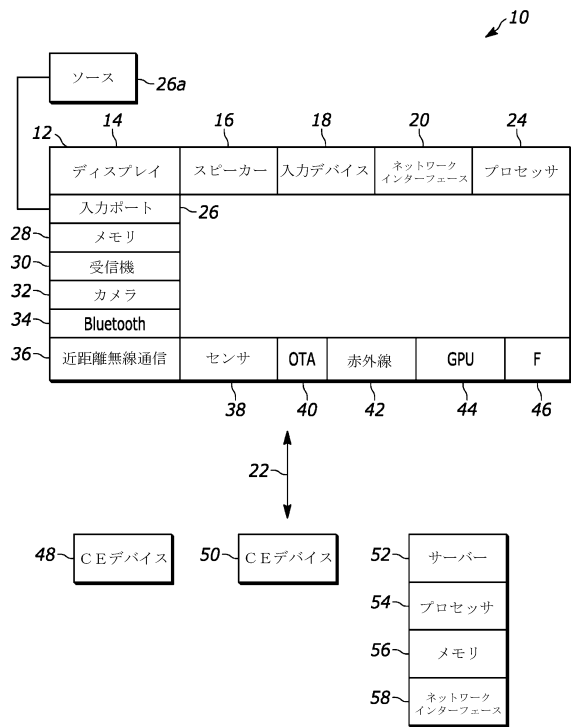
30

40

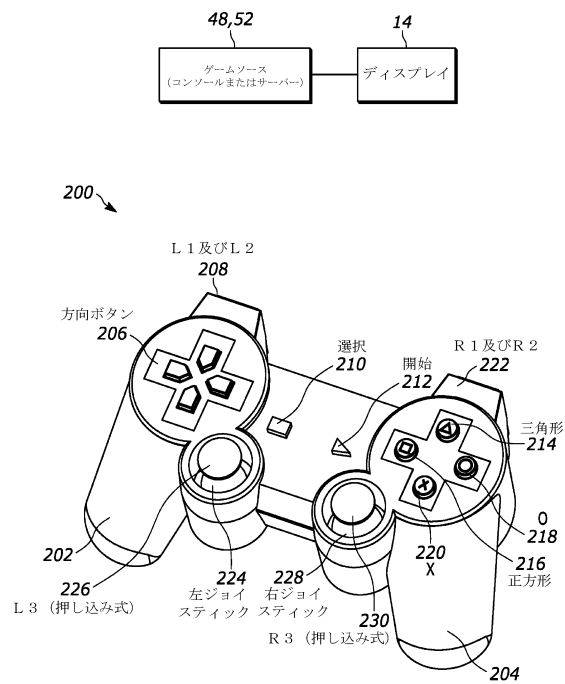
50

【図面】

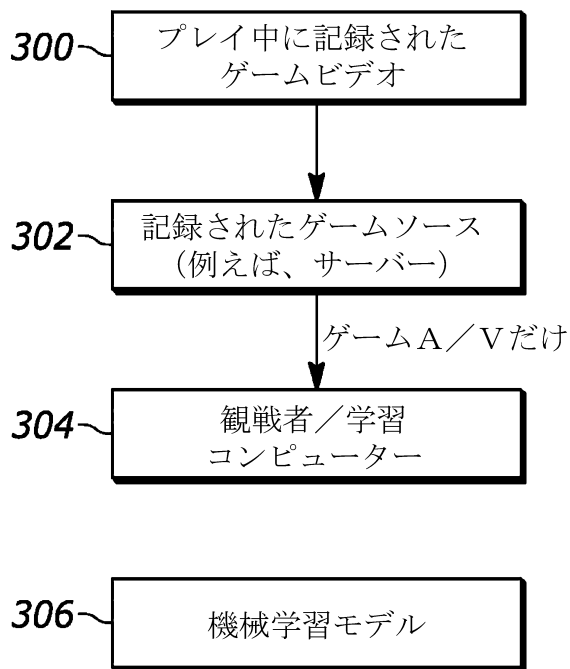
【図 1】



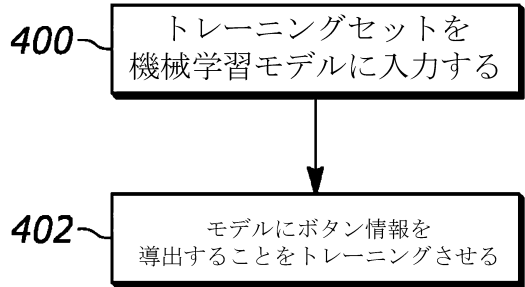
【図 2】



【図 3】



【図 4】



10

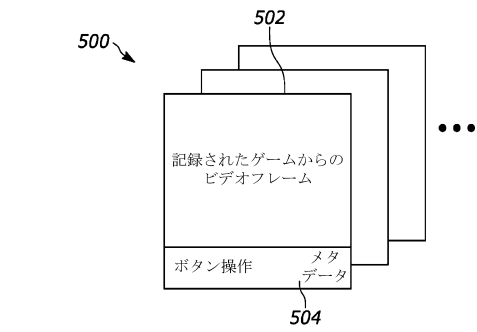
20

30

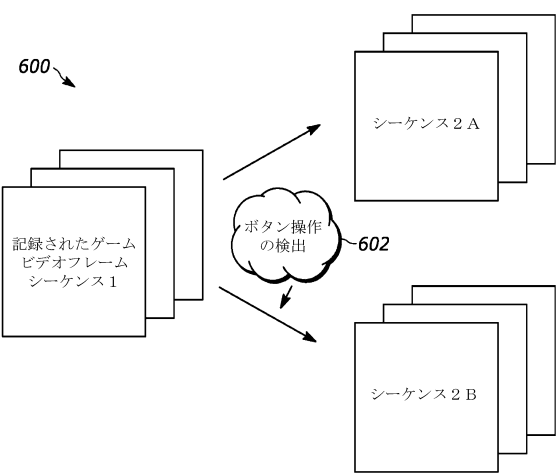
40

50

【図 5】

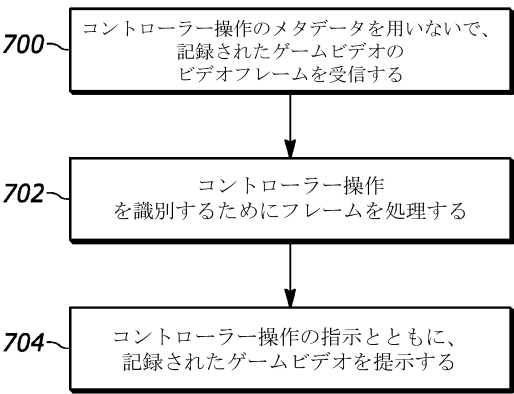


【図 6】

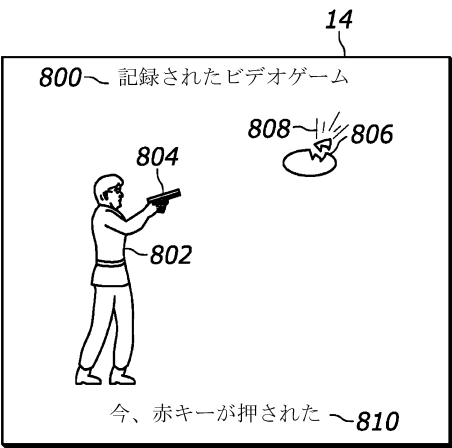


10

【図 7】



【図 8】



20

30

40

50

フロントページの続き

- (51)国際特許分類

G 0 6 T

7/00

(2017.01)

G 0 6 V

10/82

(2022.01)

F I

G 0 6 T

7/00

3 5 0 C

G 0 6 V

10/82

(56)参考文献

米国特許出願公開第 2 0 2 0 / 0 3 4 9 9 7 6 (U S , A 1)

米国特許出願公開第 2 0 2 0 / 0 3 3 8 4 4 5 (U S , A 1)

特開 2 0 1 6 - 1 1 8 8 4 9 (J P , A)

(58)調査した分野 (Int.Cl., D B 名)

A 6 3 F 1 3 / 0 0 - 9 8

A 6 3 F 9 / 2 4

G 0 6 T 7 / 0 0

G 0 6 V 1 0 / 8 2

G 0 6 N 2 0 / 0 0