

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7467615号
(P7467615)

(45)発行日 令和6年4月15日(2024.4.15)

(24)登録日 令和6年4月5日(2024.4.5)

(51)国際特許分類		F I	
H O 4 N	21/8355(2011.01)	H O 4 N	21/8355
H O 4 N	21/442(2011.01)	H O 4 N	21/442
H O 4 N	21/488(2011.01)	H O 4 N	21/488
G O 6 T	7/20(2017.01)	G O 6 T	7/20 3 0 0 Z
G O 6 T	7/60(2017.01)	G O 6 T	7/60
請求項の数 8 (全21頁) 最終頁に続く			
(21)出願番号	特願2022-521716(P2022-521716)	(73)特許権者	310021766
(86)(22)出願日	令和2年9月30日(2020.9.30)		株式会社ソニー・インタラクティブエン
(65)公表番号	特表2022-551661(P2022-551661		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4年12月12日(2022.12.12)	(74)代理人	100105924
(86)国際出願番号	PCT/US2020/053523		弁理士 森下 賢樹
(87)国際公開番号	WO2021/071721	(72)発明者	イエ、シャオヨン
(87)国際公開日	令和3年4月15日(2021.4.15)		アメリカ合衆国、カリフォルニア州 9
審査請求日	令和4年4月27日(2022.4.27)		4 4 0 4、サン マテオ、ブリッジポイ
(31)優先権主張番号	16/597,698		ント パークウェイ 2 2 0 7
(32)優先日	令和1年10月9日(2019.10.9)	(72)発明者	ベネデット、ウォーレン
(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
前置審査			4 4 0 4、サン マテオ、ブリッジポイ
		審査官	富樫 明
			最終頁に続く

(54)【発明の名称】 ブロックチェーンを使用したフェイクビデオ検出

(57)【特許請求の範囲】

【請求項1】

システムであって、

少なくとも1つのプロセッサであって、前記プロセッサに、

ビデオ情報を記憶する少なくとも1つのブロックチェーンにアクセスすることと、

前記ブロックチェーンに追加しようとする第1のビデオが、前記ブロックチェーンで表される元のビデオの改変されたコピーであるかどうかを判断することと、

前記ブロックチェーンに追加しようとする前記第1のビデオが前記元のビデオの改変されたコピーであると判断したことに応じて、前記第1のビデオに関連する情報を前記ブロックチェーンに追加することを拒否するか、または前記第1のビデオが前記元のビデオから改変されている旨の表示とともに前記第1のビデオを前記ブロックチェーンに追加することと、

を実行させることが可能な命令を格納した少なくとも1つのコンピュータストレージにアクセスするように構成された前記少なくとも1つのプロセッサ

を備え、

前記判断することは、

前記第1のビデオのフレームのシーケンスに関連する音声を受信するように構成され、スペクトルを出力するように構成された少なくとも1つの周波数変換部と、

前記スペクトルを受信し、前記音声を表す音声特徴ベクトルを出力するように構成された少なくとも1つのニューラルネットワークと、

前記音声特徴ベクトルを受信し、それに基づいて、前記音声が入元の音声から改変されていることに関する表示を出力するように構成された少なくとも1つの分析モジュールであって、自然な人間の発話特性を学習するようにトレーニングされた少なくとも1つの分析モジュールと

にアクセスすることを備える、前記システム。

【請求項2】

前記命令が、前記プロセッサに、

前記ブロックチェーンに追加しようとする前記第1のビデオが前記元のビデオの改変されたコピーであると判断したことに応じて、前記第1のビデオに関連する情報を前記ブロックチェーンに追加することを拒否すること

を実行させることが可能である、請求項1に記載のシステム。

【請求項3】

前記命令が、前記プロセッサに、

前記ブロックチェーンに追加しようとする前記第1のビデオが前記元のビデオの改変されたコピーであると判断したことに応じて、前記第1のビデオが前記元のビデオから改変されている旨の表示とともに前記第1のビデオを前記ブロックチェーンに追加すること

を実行させることが可能である、請求項1に記載のシステム。

【請求項4】

前記命令が、前記プロセッサに、

前記第1のビデオの顔の、または前記第1のビデオの前記顔と背景との間の、または両方の少なくとも1つのテクスチャの不規則性が前記第1のビデオに存在するかどうかを判断すること、及び

前記テクスチャの不規則性に基づいて、前記第1のビデオが少なくとも部分的に前記元のビデオから改変されているかどうかを判断すること

によって、前記第1のビデオが少なくとも部分的に前記元のビデオの改変されたコピーであるかどうかを判断すること

を実行させることが可能である、請求項1に記載のシステム。

【請求項5】

前記命令が、前記プロセッサに、

前記第1のビデオの少なくとも一部分を表すスペクトルを出力するための少なくとも1つの離散フーリエ変換(DFT)を使用すること、及び

前記スペクトルを使用して、前記第1のビデオが少なくとも部分的に前記元のビデオから改変されているかどうかを判断すること

によって、前記第1のビデオが少なくとも部分的に前記元のビデオの改変されたコピーであるかどうかを判断すること

を実行させることが可能である、請求項1に記載のシステム。

【請求項6】

前記命令が、前記プロセッサに、

前記スペクトルの少なくとも1つの不規則性を検出することであって、前記スペクトルの前記不規則性が、前記元のビデオの対応する領域よりも明るい明るさの少なくとも1つの領域を含む、前記検出することと、

前記スペクトルの前記不規則性を検出したことに応じて、前記第1のビデオが改変されていると判断すること

を実行させることが可能である、請求項5に記載のシステム。

【請求項7】

前記命令が、前記プロセッサに、

前記第1のビデオのフレームのシーケンスを受信することと、

前記フレームのシーケンスに示されている人の顔の動きが自然な動きを示すかどうかを表す特徴ベクトルを出力することと、

前記特徴ベクトルに基づいて、前記フレームのシーケンスが前記元のビデオから改変さ

10

20

30

40

50

れていると判断することと

を実行させることが可能である、請求項 1 に記載のシステム。

【請求項 8】

前記命令が、前記プロセッサに、

前記第 1 のビデオに対してビデオハッシュを実行することと、

前記第 1 のビデオに対する前記ビデオハッシュを前記ブロックチェーンで表される前記元のビデオのビデオハッシュと比較することと、

前記第 1 のビデオに対する前記ビデオハッシュを前記元のビデオの前記ビデオハッシュと比較したことに応えて、前記第 1 のビデオが前記元のビデオの改変されたコピーであるかどうかを判断することと、

を実行させることが可能である、請求項 1 に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、一般に、コンピュータ技術に必然的に根差し、具体的な技術的改善をもたらす、技術的に独創的で非定型的な解決策に関する。

【背景技術】

【0002】

本明細書で理解されるように、深層学習アルゴリズムと組み合わされた最新のデジタル画像処理は、ある人物のビデオ画像を別の人物の画像に改変する、または人物のビデオを、その人物の声で、その人物が実際には決して話さなかった言葉を話すように改変する、興味深くかつ面白いが、潜在的に邪悪な能力を示す。このような機能は気軽に使用できるが、その機能はまた、あたかもその人物が侮辱的なことを話しているかのように見せることで、個人の名誉を毀損するために使用される可能性がある。したがって、ビデオが本物であるのか、それとも機械学習によって生成されたフェイクであるのかを判断するための技法が本明細書に提供される。

【発明の概要】

【0003】

したがって、システムは、ビデオ情報を記憶する少なくとも 1 つのブロックチェーンにアクセスするためにプロセッサによって実行可能な命令で少なくとも 1 つのコンピュータストレージにアクセスするように構成された少なくとも 1 つのプロセッサを含む。命令はまた、ブロックチェーンに追加しようとする第 1 のビデオが、ブロックチェーンで表される元のビデオの改変されたコピーであるかどうかを判断するために実行可能である。命令はさらに、ブロックチェーンに追加しようとする第 1 のビデオが元のビデオの改変されたコピーであると判断したことに応えて、第 1 のビデオに関連する情報をブロックチェーンに追加することを拒否するか、または第 1 のビデオが元のビデオから改変されている旨の表示とともに第 1 のビデオをブロックチェーンに追加するために実行可能である。

【0004】

いくつかの例では、命令は、ブロックチェーンに追加しようとする第 1 のビデオが元のビデオの改変されたコピーであると判断したことに応えて、第 1 のビデオに関連する情報をブロックチェーンに追加することを拒否するために実行可能である。他の例では、命令は、ブロックチェーンに追加しようとする第 1 のビデオが元のビデオの改変されたコピーであると判断したことに応えて、第 1 のビデオが元のビデオから改変されている旨の表示とともに第 1 のビデオをブロックチェーンに追加するために実行可能である。

【0005】

例示的な実施態様では、命令は、第 1 のビデオ内の顔の、または第 1 のビデオの顔と背景との間の、または両方の少なくとも 1 つのテクスチャの不規則性が第 1 のビデオに存在するかどうかを判断すること、及び第 1 のビデオが、テクスチャの不規則性に基づいて元のビデオから少なくとも部分的に改変されているかどうかを判断することによって、第 1 のビデオが少なくとも部分的に元のビデオの改変されたコピーであるかどうかを判断する

10

20

30

40

50

ために実行可能であってよい。

【0006】

例示的な実施態様では、命令は、第1のビデオの少なくとも一部分を表すスペクトルを出力するための少なくとも1つの離散フーリエ変換(DFT)を使用すること、及びスペクトルを使用して、第1のビデオが少なくとも部分的に元のビデオから改変されているかどうかを判断することによって、第1のビデオが少なくとも部分的に元のビデオの改変されたコピーであるかどうかを判断するために実行可能であってよい。このような実施態様では、命令は、元のビデオの対応する領域よりも明るい明るさの少なくとも1つの領域を含み得るスペクトルの少なくとも1つの不規則性を検出するために実行可能であってよい。命令は、スペクトルの不規則性を検出したことに応じて、第1のビデオが改変されていると判断するために実行可能である可能性がある。

10

【0007】

非限定的な例では、命令は、第1のビデオのフレームのシーケンスを受信するために、及びフレームのシーケンスに示される人の顔の動きが自然な動きを示すかどうかを表す特徴ベクトルを出力するために実行可能である。命令はさらに、特徴ベクトルに基づいて、フレームのシーケンスが元のビデオから改変されているかどうかを判断するために実行可能であってよい。

【0008】

所望される場合、システムは、第1のビデオのフレームのシーケンスに関連する音声を受信するために構成され、スペクトルを出力するために構成された少なくとも1つの周波数変換を含み得る。システムはまた、スペクトルを受信し、音声を表す音声特徴ベクトルを出力するために構成された少なくとも1つのニューラルネットワークを含み得る。さらに、システムは、音声特徴ベクトルを受信し、それに基づいて、音声が元の音声から改変されていることに関する表示を出力するために構成された、自然な人間の発話特性を学習するようにトレーニングされた少なくとも1つの分析モジュールを含むことができる。

20

【0009】

非限定的な例示的な実施形態では、命令は、第1のビデオに対してビデオハッシュを実行するために実行可能であってよい。命令はさらに、第1のビデオに対するビデオハッシュをブロックチェーンで表される元のビデオのビデオハッシュと比較するために、及び第1のビデオに対するビデオハッシュをブロックチェーンで表される元のビデオのビデオハッシュと比較したことに応じて、第1のビデオが元のビデオの改変されたコピーであるかどうかを判断するために実行可能であってよい。

30

【0010】

別の態様では、方法は、元のビデオから派生した第1のビデオが元のビデオから改変されていることを検出することを含む。方法はまた、第1のビデオが改変されている旨の表示とともに、第1のビデオを、元のビデオを表す情報を含むブロックチェーンに追加すること、または第1のビデオをブロックチェーンに追加する要求を拒否することのどちらかを含む。

【0011】

別の態様では、装置は、元のビデオの少なくとも第1のフレームについて、第1のフレームに対するハッシュを作成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同様に含む少なくとも1つの記憶媒体を含む。命令はまた、第1のフレームのハッシュの要素を第1のフレームの既知の場所にステガノグラフィ的に埋め込み、元のビデオのコピーの少なくとも第2のフレームについて、第1のフレームの既知の場所に対応する第2のフレームの場所以外の第2のフレームの場所からの情報のみを使用して、第2のフレームに対するハッシュを作成するために実行可能である。命令は、第1のフレームに対するハッシュを第2のフレームに対するハッシュと比較するために、及び比較に応じて、元のビデオのコピーが元のビデオから改変されているかどうかを示すために実行可能である。

40

【0012】

50

本出願の詳細は、その構造及び動作の両方に関して、添付図面を参照して最もよく理解することができ、図面中、同様の参照番号は同様の部分を指す。

【図面の簡単な説明】

【0013】

【図1】本原理に従った例を含む例示的なシステムのブロック図である。

【図2】本物のビデオ及び本物のビデオから派生したフェイクビデオを示す概略図である。

【図3】画像処理及び周波数領域解析の両方を使用するフェイクビデオを検出するための例示的なロジックのフローチャートである。

【図4】図3で使用されているニューラルネットワークをトレーニングするための例示的なロジックのフローチャートである。

【図5】本物のビデオフレーム及び対応するフェイクビデオフレームの図を示しており、偽のフレームのアーティファクトを示している。

【図6】図3のロジックを実行するための例示的なニューラルネットワークアーキテクチャのブロック図である。

【図7】ビデオシーケンス分析を使用してフェイクビデオを検出するための例示的なロジックのフローチャートである。

【図8】図7のロジックを実行するための例示的なニューラルネットワークアーキテクチャのブロック図である。

【図9】フェイクビデオ生成に対処するためにブロックチェーン技術を使用するための例示的なロジックのフローチャートである。

【図10】インターネットサービスプロバイダ（ISP）／配信業者がビデオを公の場から削除できるように、ISPまたは配信者にフェイクビデオを報告するための例示的なユーザーインターフェース（UI）のスクリーンショットである。

【図11】ビデオに埋め込まれた検証ハッシュとともにビデオを記録、またはアップロード、またはダウンロードするための例示的なロジックのフローチャートである。

【図12】真正性の検証のためにハッシュが使用される図11で記録またはアクセスされたビデオを再生するための例示的なロジックのフローチャートである。

【図13】以前の原理を使用したハイブリッドロジックの例示的なロジックのフローチャートである。

【図14】2セットの本物の画像と改変された画像の例示的な照明アーティファクトを示す。

【図15】画像内の例示的な敵対的生成ネットワーク（GAN）アーティファクトまたは不規則性を示す。

【図16】本物の画像及び改変された画像を使用して、別のGAN関連のアーティファクトまたは不規則性を示す。

【発明を実施するための形態】

【0014】

本開示は、概して、限定するものではないが、コンピュータゲームネットワーク及びスタンドアロンコンピュータシミュレーションシステム等のコンピュータシミュレーションネットワーク等の家電製品（CE）デバイスの態様を含むコンピュータエコシステムに関する。本明細書のシステムは、クライアントコンポーネントとサーバコンポーネントとの間でデータが交換され得るようにネットワークを通じて接続されたサーバ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る。クライアントコンポーネントは、Sony Play Station（登録商標）等のゲームコンソールまたはMicrosoftもしくはNintendoもしくは他の製造者によって作成されたゲームコンソール、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等のポータブルコンピュータ、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つまたは複数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

10

20

30

40

50

クライアントコンピュータの一部は、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 ComputerもしくはGoogle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。これらの動作環境は、下記に説明されるインターネットサーバによってホストされるウェブサイトにアクセスできる、MicrosoftもしくはGoogleもしくはMozillaによって作成されたブラウザ、または他のブラウザプログラム等、1つまたは複数のブラウジングプログラムを実行するために使用され得る。また、本原理に従った動作環境を使用して、1つまたは複数のコンピュータゲームプログラムを実行し得る。

【0015】

サーバ及び／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介してデータを受信及び送信するサーバを構成する命令を実行する1つまたは複数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。サーバまたはコントローラは、Sony Play Station（登録商標）等のゲームコンソール、パーソナルコンピュータ等によってインスタンス化され得る。

【0016】

クライアントとサーバとの間でネットワークを通じて情報を交換し得る。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つまたは複数のサーバは、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てよい。

【0017】

本明細書で使用される場合、命令は、システムにおいて情報を処理するためにコンピュータにより実施されるステップを指す。命令は、ソフトウェア、ファームウェア、またはハードウェアで実施され、システムのコンポーネントにより実行される任意の種類のプログラム化されたステップを含み得る。

【0018】

プロセッサは、アドレス線、データ線、及び制御線等の各種の線、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、従来の任意の汎用シングルチッププロセッサまたは汎用マルチチッププロセッサであり得る。

【0019】

本明細書でフローチャート及びユーザーインタフェースによって記述されるソフトウェアモジュールは、様々なサブルーチン、手順等を含み得る。本開示を限定することなく、特定のモジュールによって実行されるように規定されたロジックは、他のソフトウェアモジュールに再分配すること、及び／または単一のモジュールにまとめて集約すること、及び／または共有可能ライブラリで利用できるようにすることが可能である。

【0020】

本明細書に説明される本原理は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの組み合わせとして実装することができる。したがって、実例となるコンポーネント、ブロック、モジュール、回路、及びステップは、それらの機能の観点から説明される。

【0021】

さらに上記に指摘したものについて、下記に説明される論理ブロック、モジュール、及び回路は、汎用プロセッサ、デジタルシグナルプロセッサ（DSP）、フィールドプログラマブルゲートアレイ（FPGA）もしくは特定用途向け集積回路（ASIC）等の他のプログラマブル論理デバイス、個別ゲートもしくはトランジスタ論理、個別ハードウェアコンポーネント、または本明細書に説明される機能を行うように設計されたそれらのいずれかの組み合わせによって実装できるまたは行うことができる。プロセッサは、コントロ

10

20

30

40

50

ーラもしくは状態機械、またはコンピューティングデバイスの組み合わせによって実装することができる。

【0022】

以下で説明される機能及び方法は、ソフトウェアにおいて実装されるとき、限定ではないが、Java（登録商標）、C#、またはC++等の適切な言語において記述することができ、ランダムアクセスメモリ（RAM）、リードオンリメモリ（ROM）、電氣的消去可能プログラマブルリードオンリメモリ（EEPROM）、コンパクトディスクリードオンリメモリ（CD-ROM）、またはデジタル多用途ディスク（DVD）等の他の光ディスクストレージ、磁気ディスクストレージもしくは着脱可能サムドライブ等を含む他の磁気記憶装置等のコンピュータ可読記憶媒体に記憶することができ、またはそれらを通して伝送することができる。接続は、コンピュータ可読媒体を確立し得る。このような接続は、例として、光ファイバ、同軸ワイヤ、デジタル加入者回線（DSL）、及びツイストペアワイヤを含む有線ケーブルを含み得る。このような接続には、赤外線及び無線を含む無線通信接続が含まれ得る。

10

【0023】

ある実施形態に含まれるコンポーネントを、他の実施形態において任意の適切な組み合わせで使用することができる。例えば、本明細書で説明される、及び／または図に描かれる様々なコンポーネントのいずれも、組み合わされてもよく、交換されてもよく、または他の実施形態から除外されてもよい。

【0024】

「A、B、及びCの少なくとも1つを有するシステム」（同様に「A、B、またはC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を単独で、Bを単独で、Cを単独で、A及びBと一緒に、A及びCと一緒に、B及びCと一緒に、及び／またはA、B、及びCと一緒に有するシステム等を含む。

20

【0025】

ここで具体的に図1を参照すると、例示的なシステム10が示されており、このシステムは、本原理による、上で述べられかつ以下に詳述される、例示的なデバイスの1つまたは複数を含み得る。システム10に含まれる例示的なデバイスのうちの第1のデバイスは、TVチューナ（同等にTVを制御するセットトップボックス）を備えたインターネット対応TV等の、ただしこれに限定されないオーディオビデオデバイス（AVD）12といった家電製品（CE）デバイスである。しかし、AVD12は、代わりに、電気器具または家庭用品、例えば、コンピュータ制御でインターネット対応の冷蔵庫、洗濯機、または乾燥機であってよい。代わりに、AVD12は、また、コンピュータ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、ノートブックコンピュータ、例えば、コンピュータ制御型インターネット対応時計、コンピュータ制御型インターネット対応ブレスレット、他のコンピュータ制御型インターネット対応デバイス等のウェアラブルコンピュータ制御デバイス、コンピュータ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制御型で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ってよい。それにも関わらず、AVD12は、本原理を実施する（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）ように構成されることを理解されたい。

30

【0026】

したがって、こ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つまたは複数のディスプレイ14を含むことができ、ディスプレイ14は、高解像度または「4K」もしくはそれ以上の超高解像度のフラットスクリーンによって実装されてよく、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ってよい。AVD12は、本原理に従って音声を出力するための1つまたは複数のスピーカ16、及び、例え

40

50

ば、A V D 1 2 を制御するために可聴コマンドを A V D 1 2 に入力するための、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 1 8 を含み得る。例示的な A D V 1 2 はまた、1 つまたは複数のプロセッサ 2 4 等の制御の下、インターネット、W A N、L A N 等の少なくとも1つのネットワーク 2 2 を通じて通信するための1 つまたは複数のネットワークインタフェース 2 0 を含み得る。また、グラフィックプロセッサ 2 4 A を含み得る。したがって、インタフェース 2 0 は、限定ではないが、W i - F i (登録商標) 送受信機であり得、W i - F i (登録商標) 送受信機は、限定するものではないが、メッシュネットワーク送受信機等の無線コンピュータネットワークインタフェースの例である。プロセッサ 2 4 は、例えば、ディスプレイ 1 4 を、画像を提示するように制御することや、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 A D V 1 2 の他の要素を含む、本原理を実施するように A D V 1 2 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フェース 2 0 は、例えば、有線もしくは無線のモデムもしくはルータ、または、例えば、無線テレフォニ送受信機もしくは上述した W i - F i (登録商標) 送受信機等の他の適切なインタフェースであってよいことに留意されたい。

【0027】

上記に加えて、A D V 1 2 はまた、例えば、別の C E デバイスに（例えば、有線接続を使用して）物理的に接続する高解像度マルチメディアインタフェース（H D M I (登録商標)）ポートもしくは U S B ポート、及び／またはヘッドフォンを通して A D V 1 2 からユーザーに音声を提示するために A V D 1 2 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 つまたは複数の入力ポート 2 6 を含み得る。例えば、入力ポート 2 6 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 2 6 a に有線または無線で接続されてよい。したがって、ソース 2 6 a は、例えば、別個のもしくは統合されたセットトップボックス、または衛星受信機であってよい。または、ソース 2 6 a は、以下に詳述されるチャンネル割り当て目的でユーザーが好みと見なし得るコンテンツを含むゲームコンソールまたはディスクプレイヤであってよい。ソース 2 6 a は、ゲームコンソールとして実装されるとき、C E デバイス 4 4 に関連して以下で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。

【0028】

A D V 1 2 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 つまたは複数のコンピュータメモリ 2 8 を含んでよく、これらのメモリは、場合によっては、スタンドアロンデバイスとして A D V のシャーシ内で、または A V プログラムを再生するために A V D 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（P V R）もしくはビデオディスクプレイヤとして、または取り外し可能メモリ媒体として具現化されてよい。また、いくつかの実施形態では、A V D 1 2 は、限定するものではないが、少なくとも1つの衛星もしくは携帯電話の中継塔から地理的位置情報を受信し、情報をプロセッサ 2 4 に提供し、及び／または A V D 1 2 がプロセッサ 2 4 と併せて配置される高度を判断するように構成される、携帯電話受信機、G P S 受信機、及び／または高度計 3 0 等の位置受信機または場所受信機を含むことができる。しかしながら、例えば、3 つの次元全てにおいて、例えば、A D V 1 2 の場所を決定するために、本原理に従って、携帯電話受信機、G P S 受信機、及び／または高度計以外の別の適切な位置受信機が使用され得ることを理解されたい。

【0029】

A D V 1 2 の説明を続けると、いくつかの実施形態では、A D V 1 2 は、例えば、熱探知カメラ、ウェブカメラ等のデジタルカメラ、及び／または A D V 1 2 に組み込まれ、本原理に従って写真／画像及び／またはビデオを収集するようにプロセッサ 2 4 によって制御可能なカメラであってもよい1 つまたは複数のカメラ 3 2 を含み得る。また、A V D 1 2 に含まれるのは、B l u e t o o t h (登録商標) 及び／または近距離無線通信（N F C）技術を各々使用して、他のデバイスと通信するための B l u e t o o t h (登録商標) 送受信機 3 4 及び他の N F C 要素 3 6 であってよい。例示的な N F C 要素は、無線周波数識別（R F I D）要素であってもよい。

10

20

30

40

50

【0030】

さらにまた、A V D 1 2 は、プロセッサ 2 4 に入力を提供する 1 つまたは複数の補助センサ 3 7（例えば、加速度計、ジャイロ스코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（I R）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）等）を含み得る。A V D 1 2 は、プロセッサ 2 4 への入力を提供する O T A（無線）T V 放送を受信するための無線 T V 放送ポート 3 8 を含み得る。上記に加えて、A V D 1 2 はまた、赤外線データ協会（I R D A）デバイス等の赤外線（I R）送信機及び／または I R 受信機及び／または I R 送受信機 4 2 を含み得ることに留意されたい。A V D 1 2 に給電するためのバッテリー（図示せず）が備えられてよい。

10

【0031】

さらに図 1 を参照すると、A D V 1 2 に加えて、システム 1 0 は、1 つまたは複数の他の C E デバイスタイプを含んでもよい。一例では、第 1 の C E デバイス 4 4 は、A D V 1 2 に直接送信されるコマンドを介して及び／または後述のサーバを通じて、コンピュータゲームの音声及びビデオを A D V 1 2 に送信するために使用され得る一方、第 2 の C E デバイス 4 6 は第 1 の C E デバイス 4 4 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 C E デバイス 4 6 は、示されたようにプレーヤ 4 7 によって装着される V R ヘッドセットとして構成され得る。示される例では、2 つの C E デバイス 4 4、4 6 のみが示されるが、より少ないまたはより多くのデバイスが使用されてよいことを理解されたい。例えば、以下の原理は、ゲームコンソールによって 1 つまたは複数の A V D 1 2 に供給されるコンピュータゲームのプレイ中に互いと通信するそれぞれのヘッドセットをした複数のプレーヤ 4 7 を説明する。

20

【0032】

示される例では、本原理を示すために、3 つのデバイス 1 2、4 4、4 6 は全て、例えば家庭のエンターテインメントネットワークのメンバーであること、または少なくとも、住宅等の場所において相互に近接して存在していることが想定される。しかし、本原理は、他に明確に主張されない限り、破線 4 8 によって示される特定の場所に限定されない。

【0033】

例示的な非限定的な第 1 の C E デバイス 4 4 は、上述のデバイス、例えば、ポータブル無線ラップトップコンピュータまたはノートブックコンピュータまたはゲームコントローラのいずれか 1 つによって確立されてもよく、したがって、以下で説明される 1 つまたは複数のコンポーネントを有してもよい。第 1 の C E デバイス 4 4 は、例えば、A V 再生コマンド及び一時停止コマンドを A V D 1 2 に発行するためのリモコン装置（R C）であってよく、またはタブレットコンピュータ、有線もしくは無線リンクを介して A V D 1 2 及び／またはゲームコンソールと通信するゲームコントローラ、パーソナルコンピュータ、無線電話等のより高性能のデバイスであってよい。

30

【0034】

したがって、第 1 の C E デバイス 4 4 は、ディスプレイをタッチしてユーザー入力信号を受信するためにタッチ対応であってよい 1 つまたは複数のディスプレイ 5 0 を含み得る。第 1 の C E デバイス 4 4 は、本原理に従って音声を出力するための 1 つまたは複数のスピーカ 5 2、及び、例えば、デバイス 4 4 を制御するために可聴コマンドを第 1 の C E デバイス 4 4 に入力するための、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 5 4 を含み得る。例示的な第 1 の C E デバイス 4 4 はまた、1 つまたは複数の C E デバイスプロセッサ 5 8 の制御の下、ネットワーク 2 2 を通じて通信するための 1 つまたは複数のネットワークインタフェース 5 6 を含み得る。また、グラフィックプロセッサ 5 8 A を含み得る。したがって、インタフェース 5 6 は、限定ではないが、W i - F i（登録商標）送受信機であり得、W i - F i（登録商標）送受信機は、メッシュネットワークインタフェースを含む無線コンピュータネットワークインタフェースの例である。プロセッサ 5 8 は、例えば、ディスプレイ 5 0 を、画像を提示するように制御することや、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第 1 の C E デバイス 4 4 の

40

50

他の要素を含む、本原理を実施するように第1のCEデバイス44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フェース56は、例えば、有線もしくは無線のモデムもしくはルータ、または、例えば、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フェースであってよいことに留意されたい。

【0035】

上記に加えて、第1のCEデバイス44はまた、例えば、別のCEデバイスに（例えば、有線接続を使用して）物理的に接続する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通して第1のCEデバイス44からユーザーに音声を提示するために第1のCEデバイス44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つまたは複数の入力ポート60を含み得る。第1のCEデバイス44は、さらに、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つまたは複数の有形コンピュータ可読記憶媒体62を含み得る。また、いくつかの実施形態では、第1のCEデバイス44は、限定するものではないが、例えば、三角測量を使用して、少なくとも1つの衛星及び／または携帯電話の基地局から地理的位置情報を受信し、情報をCEデバイスプロセッサ58に提供し、及び／または第1のCEデバイス44がCEデバイスプロセッサ58と併せて配置される高度を判断するように構成される、携帯電話受信機及び／またはGPS受信機、及び／または高度計64等の位置受信機または場所受信機を含むことができる。しかしながら、例えば、3つの次元全てにおいて、例えば、第1のCEデバイス44の場所を決定するために、本原理に従って、携帯電話及び／またはGPS受信機及び／または高度計以外の別の適切な位置受信機が使用され得ることを理解されたい。

【0036】

第1のCEデバイス44の説明を続けると、いくつかの実施形態では、第1のCEデバイス44は、例えば、熱探知カメラ、ウェブカメラ等のデジタルカメラ、及び／または第1のCEデバイス44に組み込まれ、本原理に従って写真／画像及び／またはビデオを収集するようにCEデバイスプロセッサ58によって制御可能なカメラであってもよい1つまたは複数のカメラ66を含み得る。また、第1のCEデバイス44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68及び他のNFC要素70であってよい。例示のNFC要素は、無線周波数識別（RFID）要素であってもよい。

【0037】

さらにまた、第1のCEデバイス44は、CEデバイスプロセッサ58に入力を提供する1つまたは複数の補助センサ72（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）等）を含み得る。第1のCEデバイス44は、CEデバイスプロセッサ58に入力を提供する、例えば、1つまたは複数の気候センサ74（例えば、気圧計、湿度センサ、風センサ、光センサ、温度センサ等）及び／または1つまたは複数の生体認証センサ76等のさらに他のセンサを含み得る。上記に加えて、いくつかの実施形態では、第1のCEデバイス44は、赤外線（IR）送信機及び／またはIR受信機、及び／または赤外線データアソシエーション（IRDA）デバイス等のIR送受信機78も含み得ることに留意されたい。CEデバイス44に給電するためのバッテリー（図示せず）が備えられてよい。CEデバイス44は、前述の通信モード及び関連コンポーネントのうちのいずれかを介して、AVD12と通信し得る。

【0038】

第2のCEデバイス46は、CEデバイス44について示したコンポーネントの一部または全てを含み得る。CEデバイスの一方または両方のいずれかは、1つまたは複数のバッテリーによって供电され得る。

【0039】

10

20

30

40

50

ここで、上述の少なくとも1つのサーバ80を参照すると、サーバ80は、少なくとも1つのサーバプロセッサ82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可読記憶媒体84と、サーバプロセッサ82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フェース86を含む。ネットワークインタフェース86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フェースであり得ることに留意されたい。

【0040】

したがって、いくつかの実施形態では、サーバ80は、インターネットサーバまたはサーバ「ファーム」全体であってもよく、「クラウド」機能を含んでもよく、「クラウド」機能を実行してもよく、その結果、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ついての例示的な実施形態においてサーバ80を介して「クラウド」環境にアクセスすることができる。または、サーバ80は、図1に示す他のデバイスと同じ部屋かまたはその近くにある1つまたは複数のゲームコンソールまたは他のコンピュータによって実装されてもよい。

【0041】

本明細書における方法は、プロセッサ、適切に構成された特定用途向け集積回路（ASIC）もしくはフィールドプログラマブルゲートアレイ（FPGA）モジュール、または当業者によって認識される任意の他の便利な方式によって実行されるソフトウェア命令として実装されてよい。使用される場合、ソフトウェア命令は、CD ROMまたはフラッシュドライブ等の非一時的デバイスで具現化されてよい。代わりに、ソフトウェアコード命令は、無線信号もしくは光信号等の一時的構成で、またはインターネットを通じたダウンロードを介して具現化されてよい。

【0042】

ここで図2～図6を参照すると、画像が「フェイク」であるかどうか、つまり元の画像からデジタル的に改変されているかどうかを判断するための第1の技法が示されている。図2では、可聴の言葉206を話している、顔204を有する人の、ディスプレイ202に提示され得る元の画像200が示されている。画像200は、ビデオストリームからのエフェーム等の画像であってよく、ビデオストリームのフレームの一部または全ては、本明細書に開示されるように処理され得る。

【0043】

限定するものではないが、キーボード等の入力デバイス210を用いてコンピュータ208を操作する人は、画像及び／または音声を改変して、改変された可聴の言葉214を話していると描かれる可能性のある人の改変された画像212を生成し得る。本原理は、改変された画像212が実際にオリジナル200から改変されていることを検出することを対象とする。

【0044】

図3は、第1の技法で実行され得るロジックを示し、一方、図6は、図3のロジックを具現化し得る例示的なアーキテクチャを示す。ブロック300で開始すると、画像が受信される。画像は、畳み込みNN（CNN）等の第1のニューラルネットワーク（NN）を介して画像を処理することによって、ブロック302で直接分析することができる。第1のNNは、ブロック304で画像を表す特徴ベクトルを出力する。

【0045】

また、ブロック306で、画像は、顔及び／または画像の背景の、本明細書では不規則性とも呼ぶアーティファクト、及び画像内の照明の不規則性がないか分析するために顔認識モジュールに入力され得る。1つまたは複数のNNを使用し得る顔認識モジュールは、特徴ベクトルをブロック304に出力する。

【0046】

10

20

30

40

50

例として、画像（空間領域）内の顔の不規則性は、チェッカーボードのような外観を有する小さい領域を含む場合があり、デジタル改変に起因する解像度のぼやけを示す。

【0047】

さらに、画像は、例えば、周波数領域内の画像の不規則性を検出するためにCNN等の別のNNを用いてブロック310で分析される周波数スペクトルを出力する離散フーリエ変換（DFT）を使用して、ブロック308で周波数流域に変換され得る。スペクトルを表す特徴ベクトルは、ブロック304に提供される。

【0048】

例として、周波数領域の不規則性は、周波数領域内の画像のグラフィック表現の周辺に沿った1つまたは複数の輝点を含み得る。

【0049】

決定ひし形312に移動すると、1つまたは複数のNNを含む場合がある検出モジュールが、ブロック304からの特徴ベクトルを分析して、1つまたは複数の不規則性が空間領域及び／または周波数領域に存在するかどうかを判断する。不規則性が存在しない場合、プロセスは状態314で終了してよいが、いくつかの実施態様では、任意の領域に任意の不規則性が存在する場合、ブロック316で、画像がフェイクである旨の表示が返されてよい。他の実施態様では、ブロック316で、不規則性が空間領域内に存在する、及び不規則性が周波数領域に存在する両方の場合にだけ、画像がフェイクである旨の表示が返されてよい。

【0050】

図4を簡略に参照すると、本明細書に説明するNNをトレーニングするためのプロセスが示されている。ブロック400で開始すると、グラウンドトゥルースの元の改変されていない画像がNNに入力される。また、グラウンドトゥルースの改変された、つまりフェイクの画像もブロック402でNNに入力される。フェイク画像は、設計者によって、グラウンドトゥルースの元の画像から「ディープフェイク」技術を使用して生成され得る。NNは、例えば、周波数領域と空間領域の両方について上述した任意のまたは例示的な不規則性を使用して分析を開始するようにプログラムされ得る。NNは、ブロック404でグラウンドトゥルース入力に関してトレーニングされる。その後、ブロック404でNNのトレーニングを改良するために強化学習が適用され得る。

【0051】

図5は、例示的な空間領域及び周波数領域の不規則性を示す。元の画像500は、元の空間領域502及び元の周波数領域504に示されている。元の画像500の改変された画像506は、510に示される改変された空間領域画像508及び改変された周波数領域を有する。

【0052】

示されるように、改変された空間領域画像508内の領域512は、514で拡大され、示されるチェッカーボードパターンを有する。元の画像と改変された画像との間に、照明不規則性が存在する場合もある。

【0053】

1つまたは複数の周波数領域の不規則性516はまた、周波数領域510内の画像の表現において検出され得る。周波数領域の不規則性516は、示されるように、周波数領域のグラフィック描写の端縁または周辺に沿って輝点を含む場合がある。示されている例では、側面ごとに2つの輝点が存在し、周波数領域における画像改変により生じた不規則性を示している。

【0054】

図6は、図3のロジックを具現化するために使用され得る例示的なアーキテクチャを示している。改変についてテストされる画像600は、モジュール602のニューラルネットワーク（NN）604で画像の照明の不規則性、及び606で顔の解像度／不規則性チェックを検出するために空間領域の画像を分析する顔検出モジュール602に入力される。顔検出モジュール602は、画像認識原理を使用し得、1つまたは複数のNNによって

10

20

30

40

50

具現化され得る。

【0055】

また、画像600は、追加の規則を使用した直接的な分析のために、CNNである場合があるNN608に直接入力され得る。NN608が、画像の特徴ベクトルを抽出することに留意されたい。また、NN604は画像処理を実行し、十分なトレーニングデータがない場合に特に有利である。ただし、NN604、608は単一のNNによって実装される場合がある。

【0056】

さらに、画像600は、周波数領域で画像600を表すスペクトル612を出力する離散フーリエ変換(DFT)610によって処理される。スペクトル612は、スペクトルを分析するためにCNN614に送られる。

【0057】

顔認識モジュール602(照明の不規則性チェック604及び顔の解像度/アーティファクトチェック606を含む)、ならびにCNN608及び614は、空間領域と周波数領域の両方で画像600を表す特徴ベクトルのグループ616を生成する。長短期モジュール(LSTM)等のリカレントNN(RNN)等の1つまたは複数のNNによって実装され得る検出モジュール618は、特徴ベクトルを分析して、本明細書に提案される原理に従って、画像600が元の画像からのデジタル改変を含むかどうかを確認する。画像600がデジタル改変を含む場合、620で、画像600がフェイクである可能性がある旨の表示が生成される。

【0058】

図7は、改変されたビデオを検出するための第2の技法を示し、図8は、図7のロジックを具現化するための例示的なアーキテクチャを示している。ビデオクリップまたはビデオフレームの他のシーケンス等のビデオシーケンスは、ブロック700で、NNに入力される。シーケンスは、NNを使用してブロック702で分析され、NNは、704でビデオシーケンスを表す特徴ベクトルを出力する。

【0059】

ビデオシーケンスを分析する際に、NNは、例えば発話中の自然な人間の顔の動きのパターンを学習するようにトレーニングされ得る。本明細書で理解されるように、ビデオシーケンスが改変されるとき、改変者は、例えば唇の動きの自然なパターンを正確にモデル化し得ないため、わずかに不自然な動きのパターンがフェイクビデオシーケンスでNNによって検出され得る。

【0060】

また、ビデオシーケンスと関連付けられた音声は、ブロック706で周波数変換に入力される。周波数変換706によって出力されるスペクトルは、スペクトルの分析のためにブロック708でNNに提供されて、ブロック704に、音声を表す特徴ベクトルを出力する。

【0061】

付随する音声を分析する際に、NNは、ケーデンス、トーン、ピッチパターン、及び強調等の自然な人間の発話特性を学習するようにトレーニングされ得る。本明細書で理解されるように、声の音声等の音声は改変されるとき、改変者は、人間の発話の自然なパターンを正確にモデル化し得ない。したがって、不自然なケーデンスまたはトーンまたはピッチ等のわずかに不自然な発話のパターンは、フェイク音声シーケンスでNNによって検出され得る。トレーニングは図4に示す線に沿って達成され得、元のグラウンドトゥールース音声から派生したグラウンドトゥールース音声及びフェイクグラウンドトゥールース音声が、トレーニングセットとして使用される。

【0062】

特徴セット704は、特徴ベクトルを分析して、決定ひし形712で、入力されたビデオシーケンス及び/または付随する音声オリジナルから改変されているかどうかを検出するためにRNN710等のNNに提供され得る。異常または不規則性が見つからない場

10

20

30

40

50

合、プロセスは、状態 7 1 4 で終了し得るが、不規則性が検出された場合、ブロック 7 1 6 で、ビデオシーケンスが改変された可能性がある旨の表示が出力される。

【0063】

いくつかの実施形態では、音声またはビデオのどちらかでなんらかの不規則性が検出された場合、フェイクの表示がブロック 7 1 6 で出力される。他の実施形態では、音声とビデオの両方で不規則性が検出された場合にのみ、フェイクの表示がブロック 7 1 6 で出力される。

【0064】

図 8 は、図 7 のロジックを具現化するために使用できるアーキテクチャを示している。ビデオシーケンス 8 0 0 は、CNN 等の NN 8 0 2 に入力されて、そこから特徴ベクトル 8 0 4 を抽出する。また、発話 8 0 6 等の音声は、周波数領域で、そこから特徴ベクトルを抽出するために CNN 等の NN 8 1 0 によって分析される音声の表現を生成するために短時間フーリエ変換 (STFT) 等の周波数変換 8 0 8 に入力される。LSTM 等の RNN 等の NN 2 1 2 は、本明細書に説明する原理に従って特徴ベクトルを分析して、ブロック 8 1 4 で、ビデオシーケンス 8 0 0 及び音声 8 0 6 のあらゆる不規則性を検出する。状態 8 1 6 は、入力がフェイクである可能性がある旨の表示の出力を示す。

【0065】

ここで図 9 に目を向けると、ブロックチェーン技術及び／またはデジタルフィンガープリント技術を使用してフェイクビデオ生成に対処するための第 3 の技法が示されている。一般に、ビデオのハッシュ／署名は、例えば、ウェブブラウザを使用してスマートフォンまたは他の記録デバイス等の画像化デバイスに組み込まれ得る、またはハードウェアに符号化され得る。デジタルフィンガープリントは、ビデオ全体またはサブトラックのデータのビットから生成できるため、ビデオコンテンツが変更されると、フィンガープリントも変化する。デジタルフィンガープリントは、ビデオがどこで及びいつ最初に作成されたのかの場所及びタイムスタンプ等のメタデータとともに生成できる。ビデオを再配信しようとするたびに、配信業者はブロックチェーン上の元のブロックからの許可を要求する必要があり、新しい（コピーされた）ビデオの新しいブロックが拘束されるため、元のビデオ及びブロックチェーン上の任意のノードまでたどることが容易になる。ビデオを再度アップロードする前に、ビデオのフィンガープリントを元のフィンガープリントと照合して、追加しようとするビデオが操作されているかどうかを判断できる。

【0066】

例えば、ビデオウェブサイトにはビデオフィンガープリント検出器が組み込まれ得るため、ビデオがアップロード／ダウンロードされるたびに、ビデオは記録され、スタンプが付けられる。フィンガープリントが元のビデオフィンガーポイントに一致しないことに基づいて、ビデオがフェイクと分類された場合、ビデオはチェーン全体でこのようなものとして記録され得る。これはウイルス対策ソフトウェアをエミュレートするが、この場合、全てのユーザーが同時に保護される。

【0067】

ブロック 9 0 0 で開始すると、元の（「本物の」）ビデオはビデオのハッシュとともにビデオブロックチェーンに追加され、そのハッシュはデジタルフィンガープリントとして機能し得、通常、ビデオのピクセル値またはエンコード情報または他の画像に依存する値に基づいている。ビデオをコピーする要求はブロック 9 0 2 で受信される場合があり、要求はブロック 9 0 4 で許可される場合がある。

【0068】

ブロック 9 0 6 に移動すると、元のビデオのコピーである新しいビデオを追加してブロックチェーンに戻す要求が受信される場合がある。要求には、有効であるために、新しいビデオに対するハッシュ（フィンガープリント）が伴う場合がある。決定ひし形 9 0 8 に進むと、ブロックチェーンに追加しようとするビデオのハッシュは、ビデオのコピー元であった、元のビデオのハッシュと比較され、ハッシュが一致する場合、新しいビデオはブロック 9 1 0 でブロックチェーンに追加されてよい。

【0069】

他方、決定ひし形908で、ハッシュが一致しないと判断された場合、ロジックはブロック912に移動して、新しいビデオをブロックチェーンに追加することを拒否するか、またはそれを、新しいビデオがオリジナルから改変されているため、フェイクである場合がある旨の表示とともにブロックチェーンに追加し得る。所望される場合、改変されたビデオは、ブロックチェーンからアクセス不可、またはそれ以外の場合ブロック914で再生不可にされ得る。

【0070】

さらに、改変されたビデオが検出されると、ロジックはブロック916に移動して、インターネットサービスプロバイダ（ISP）または新しい改変されたビデオの他の配信業者に、ビデオが実際に元のビデオからデジタル的に改変されているため、新しい（改変された）ビデオを公けの場から削除するかどうかを決定するために調べられる必要があることを報告し得る。図10を説明する。

【0071】

示されるように、ユーザーインターフェース（UI）1000は、図9のロジックを実行する、及び／または図9のロジックを実行するデバイスからの情報を受信するデバイスのディスプレイ1002に提示されてよい。UI1000は、フェイクビデオが検出された可能性がある旨のプロンプト1004を含んでよい。UI1000はまた、ユーザーが、識別情報とともにフェイクの存在を配信業者または他の機関に報告することを可能にするためにセクタ1006を含み得る。

【0072】

図11及び図12は、追加のフィンガープリントロジックを示している。図11のブロック1100で開始すると、新しい元のビデオの作成に従って、及び／またはビデオのアップロードもしくはダウンロード時に、ビデオの少なくとも一部のフレーム、及びいくつかの実施形態では、ビデオの全てのフレームについて、ハッシュはフレーム内で実行される。ブロック1002に進むと、ハッシュは次に、その派生元のフレームに埋め込まれる。

【0073】

例では、ビデオフレームのハッシュは、それが肉眼では検出できず、ビデオフレーム全体で均一に分散され得るように、ビデオフレームにステガノグラフィ的に埋め込まれ得る。例えば、ステガノグラフィハッシュの各個のピクセルは、それがつねに固定位置にあるため、または場所がフレームのためのビデオメタデータに含まれている（あらゆるフレームが異なることを可能にする）ためのどちらかで、既知の場所にある場合がある。この場所を知っていることによって、ハッシュを表すピクセルをハッシュされるビデオから除外することが可能になる。言い換えれば、元のハッシュは、ステガノグラフィ的ではなく改変されたピクセルからのみ作成される。この場所はまた、ハッシュを表しているピクセルが、ハッシュに影響を与えるであろうように圧縮または改変されないことを保証するために、ビデオ圧縮アルゴリズムによって使用され得る。

【0074】

図12は、ビデオ再生ソフトウェアがこのプロセスを逆に示している。ブロック1200で開始すると、ステガノグラフィ的に埋め込まれたハッシュがビデオフレームから抽出される。ブロック1202に移動すると、ビデオフレームの残りのピクセルがハッシュされる。決定ひし形1204に進むと、新たなハッシュが、フレームから抽出されたハッシュと比較される。それらが一致する場合、フレームは元のソースビデオから改変されていないため、ロジックはブロック1206に移動して、このようなものを示し、所望される場合（全てのフレームまたは少なくとも閾値数のフレームが一致すると仮定して）ブロックチェーンにビデオを追加する。ハッシュが一致しない場合、ロジックはブロック1208に移動して、（例えば）改変されたフレームの周りの赤い境界線または強調表示により、表示されているビデオがオリジナルから改変されていることを示す。フレームの改変された部分にも輪郭を付けてよい。

【0075】

この同じ検証プロセスは、フェイクを検出し、積極的にフェイクが公開されるのを防ぐか、またはビデオに警告を付加するバックエンドサーバ上で実行され得る。

【0076】

任意の悪意のある当時者が任意の意味のある方法でソースビデオを改変すると、フレームは異なるようにハッシュする、及び／または埋め込まれたステガノグラフィハッシュが破損するであろう。この活動の両端に誠実な当事者がいる限り、ビデオの改変は検出可能である。

【0077】

図13は、上記の原理を組み合わせたハイブリッド技法が使用され得ることを示す。ブロック1300は、周波数領域分析と組み合わせた画像処理／ビデオシーケンシングの両方とも、ビデオのアーティファクト／不規則性を特定するために使用され得ることを示している。ブロック1302は、ビデオのアーティファクト／不規則性を特定するために、発話処理が上記技法のいずれかと組み合わせて使用され得ることをさらに示す。ブロック1304は、ビデオのアーティファクト／不規則性の識別が、元の（本物の）ビデオ、及び改変されているそのコピー（フェイク）を追跡するためにブロックチェーン技術と組み合わせられ得ることを示す。

【0078】

図14～図16は、図中で「フェイク」画像とラベルが付けられた、改変された画像に現れる可能性のあるアーティファクトまたは不規則性の追加の例を示す。図14の第1の本物の画像1400は、領域1404で、照明が、第1の本物の画像1400の対応する領域においてよりも明るく見える対応する改変された画像1402を生成するために改変されている。同様に、第2の本物の画像1406は、顔の領域1410での照明が、本物の画像1406よりも明るく見える改変された画像1408を生成するために改変されている。改変された画像1402、1408の解像度はまた、対応する本物の画像1400、1406の解像度よりも低く、NNが、照明の不規則性と解像度の減少のどちらかまたは両方に基づいて改変された画像を区別することを学習できることを意味する。

【0079】

図15は、小さい領域1502で、改変された画像1500を生成するために実行された敵対的生成ネットワーク（GAN）アップサンプリングのために、画像の不規則性またはアーティファクトが存在する、改変された画像1500を示す。領域1502の分解図1504に示されるように、GAN不規則性は、均一ではない無地の主題（示されている例では、様々な度合いの陰影のある草）が元の画像に現れる、画像の領域の均一に無地の色を含み得る。

【0080】

図16は、本物の画像1600、及び別の人の顔を、本物の画像1600内の対象の顔に重ね合わせることによって本物の画像1600から派生した改変された画像1602を示す。1604に示されるように、この重ね合わせによって、頭部または体の残りの部分に対する顔のずれ、この場合、頭部が示されている角度に対する鼻のずれが生じる。

【0081】

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たが、これらは限定することを意図しておらず、各種の代替的な構成が本明細書で特許請求される主題を実施するために使用されてよいことを理解されたい。

10

20

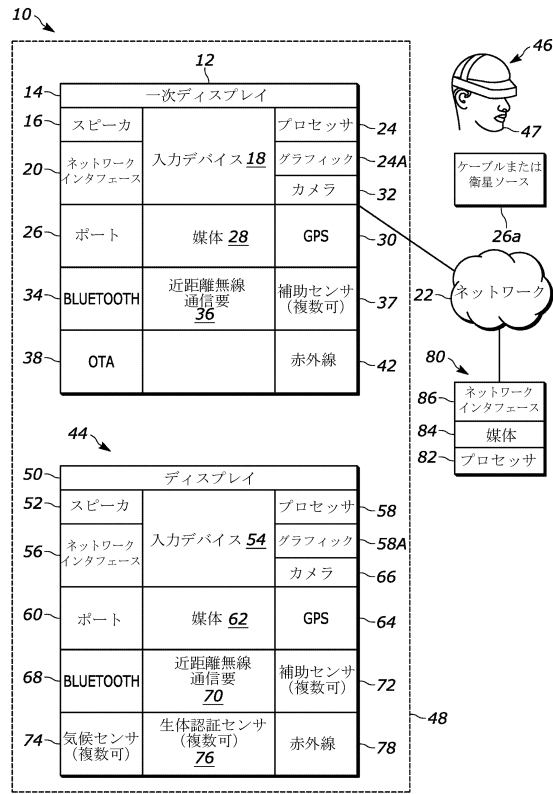
30

40

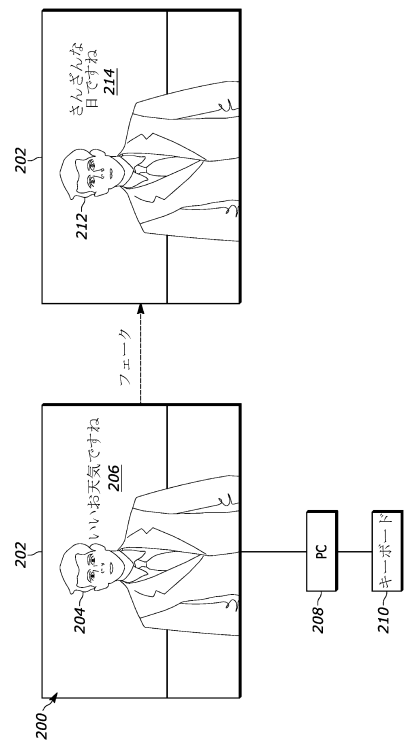
50

【図面】

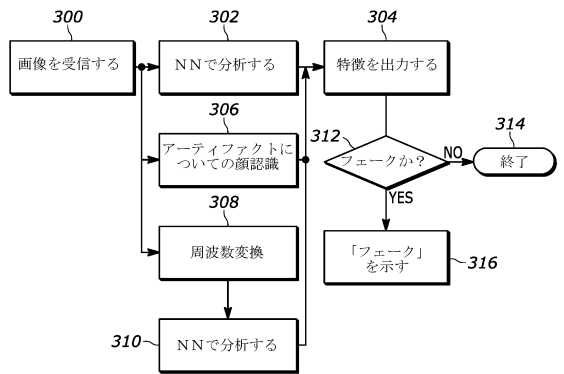
【図 1】



【図 2】



【図 3】



【図 4】



10

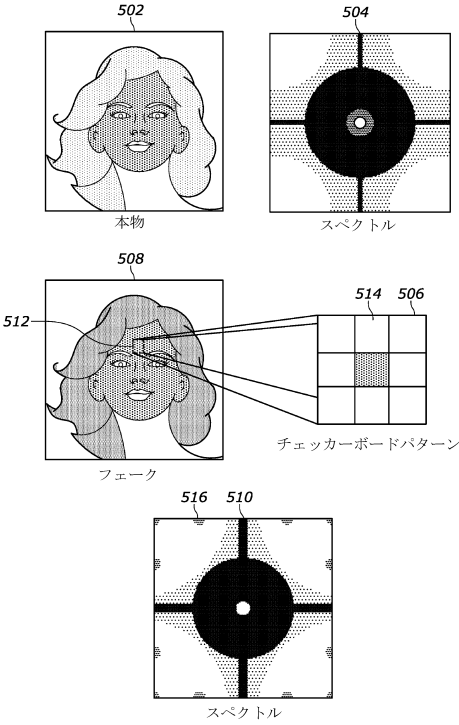
20

30

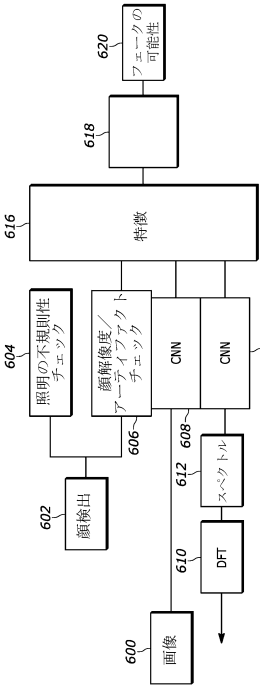
40

50

【図 5】



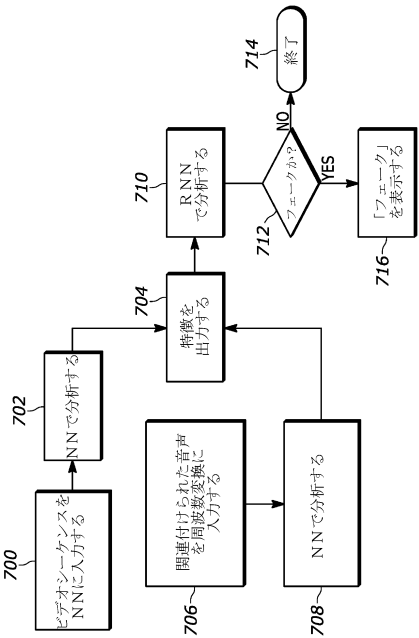
【図 6】



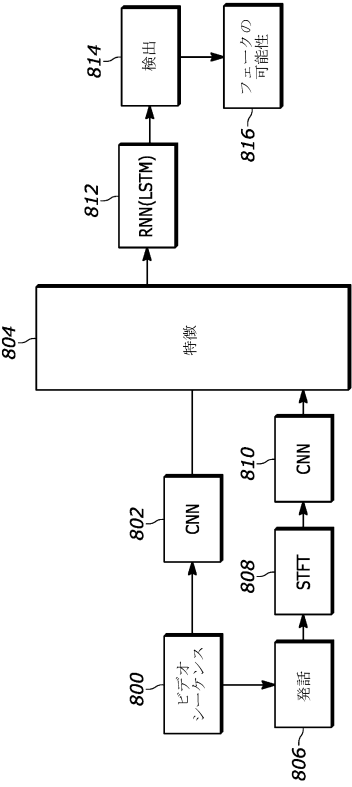
10

20

【図 7】



【図 8】

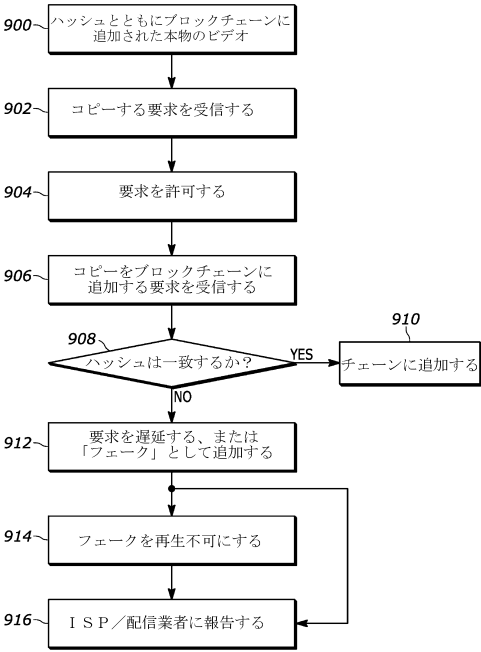


30

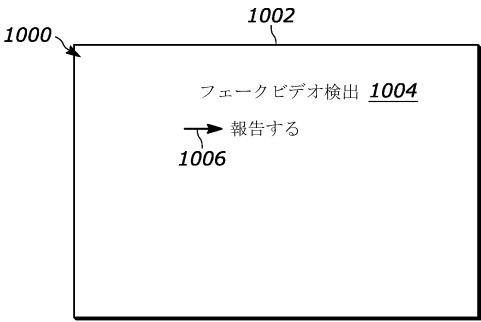
40

50

【図 9】

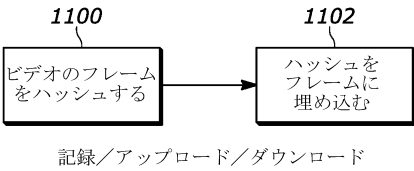


【図 10】

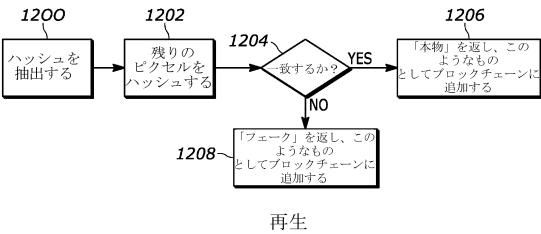


10

【図 11】



【図 12】



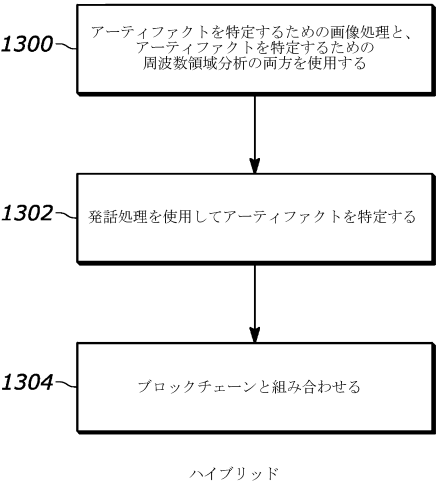
20

30

40

50

【図 1 3】



【図 1 4】

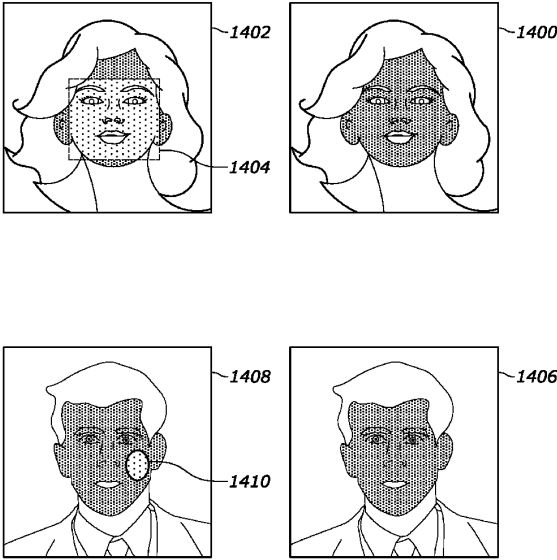
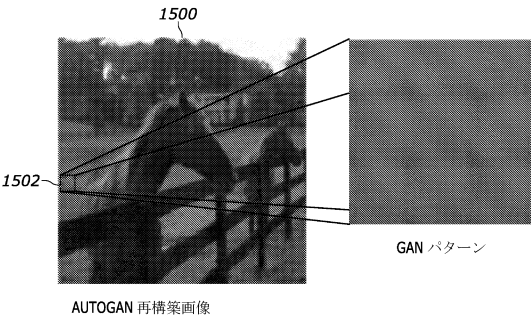


FIG. 14

【図 1 5】



【図 1 6】

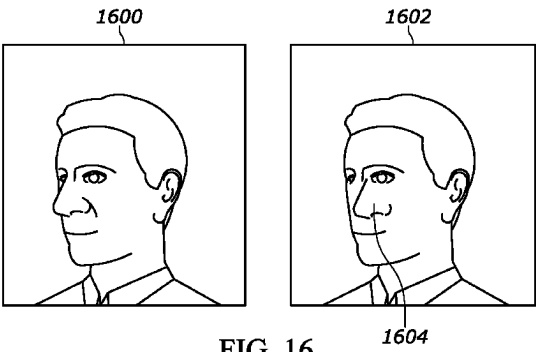


FIG. 16

フロントページの続き

(51)国際特許分類 F I
H 0 4 N 21/4627(2011.01) H 0 4 N 21/4627

(56)参考文献 特開 2 0 1 9 - 1 6 4 5 9 1 (J P , A)
特開 2 0 1 5 - 0 0 7 9 1 6 (J P , A)
特開 2 0 0 4 - 1 9 3 9 4 2 (J P , A)
国際公開第 2 0 0 7 / 0 4 6 2 8 9 (W O , A 1)
特開平 1 1 - 1 3 6 6 1 8 (J P , A)
中国特許出願公開第 1 0 8 8 0 6 7 1 8 (C N , A)
特開 2 0 0 8 - 3 0 1 5 0 0 (J P , A)
米国特許第 1 0 3 6 1 8 6 6 (U S , B 1)

(58)調査した分野 (Int.Cl., D B 名)
H 0 4 N 2 1 / 0 0 - 2 1 / 8 5 8
G 0 6 T 7 / 2 0
G 0 6 T 7 / 6 0