(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7297989号
(P7297989)

(45)発行日　令和5年6月26日(2023.6.26)

(24)登録日　令和5年6月16日(2023.6.16)

(51)国際特許分類
G06V 40/40 (2022.01)
G06T 7/20 (2017.01)

FI
G06V 40/40
G06T 7/20 300B

請求項の数 17 （全20頁）

(21)出願番号　特願2022-110618(P2022-110618)
(22)出願日　令和4年7月8日(2022.7.8)
(65)公開番号　特開2022-133463(P2022-133463A)
(43)公開日　令和4年9月13日(2022.9.13)
審査請求日　令和4年7月8日(2022.7.8)
(31)優先権主張番号　202110775389.7
(32)優先日　令和3年7月8日(2021.7.8)
(33)優先権主張国・地域又は機関
中国(CN)

(73)特許権者　514322098
ベイジン バイドゥ ネットコム サイエンス テクノロジー カンパニー リミテッド
Ｂｅｉｊｉｎｇ Ｂａｉｄｕ Ｎｅｔｃｏｍ Ｓｃｉｅｎｃｅ Ｔｅｃｈｎｏｌｏｇｙ Ｃｏ．，Ｌｔｄ．
中華人民共和国 ペキン １０００８５，ハイディアン ディストリクト，シャンディ テンス ストリート，１０番，バイドゥ キャンパス ２階
２／Ｆ Ｂａｉｄｕ Ｃａｍｐｕｓ，Ｎｏ．１０，Ｓｈａｎｇｄｉ １０ｔｈ Ｓｔｒｅｅｔ，Ｈａｉｄｉａｎ Ｄｉｓｔｒｉｃｔ，Ｂｅｉｊｉｎｇ １０００

最終頁に続く

(54)【発明の名称】　顔生体検出方法、装置、電子機器及び記憶媒体

(57)【特許請求の範囲】

【請求項１】

顔生体検出方法であって、

顔ビデオにおける連続したＮフレームの画像を取得し、前記Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得するステップであって、Ｎが１より大きい整数であるステップと、

前記Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得するステップと、

前記Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得するステップと、

前記Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、前記第２の差分画像に基づいて対応する第２の検出結果を決定するステップと、

前記第１の検出結果と前記第２の検出結果とに基づいて、前記顔ビデオ内の顔の生体検出結果を決定するステップと、を含む、

ことを特徴とする顔生体検出方法。

【請求項２】

前記Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定するステップは、

前記Ｎ－１フレームの第１の差分画像に対して時空間特徴抽出を行って、前記Ｎ－１フレームの第１の差分画像の時空間特徴ベクトルを取得するステップと、

前記時空間特徴ベクトルに対して加重計算を行って、対応する第１の加重結果を取得するステップと、
前記第１の加重結果に対して正規化処理を行って、対応する第１の検出結果を取得するステップと、を含む、
ことを特徴とする請求項１に記載の方法。
【請求項３】
前記第２の差分画像に基づいて対応する第２の検出結果を決定するステップは、
前記第２の差分画像に対して特徴抽出を行って、対応する第１の特徴ベクトルを取得するステップと、
前記第１の特徴ベクトルに対して加重計算を行って、対応する第２の加重結果を取得するステップと、
前記第２の加重結果に対して正規化処理を行って、対応する第２の検出結果を取得するステップと、を含む、
ことを特徴とする請求項１に記載の方法。
【請求項４】
前記第１の検出結果と前記第２の検出結果とに基づいて、前記顔ビデオ内の顔の生体検出結果を決定するステップは、
前記第２の検出結果が予め設定された閾値より小さいか否かを判断するステップと、
前記第２の検出結果が前記予め設定された閾値より小さいことに応答して、前記第２の検出結果に基づいて前記顔ビデオ内の顔の生体検出結果を決定するステップと、
前記第２の検出結果が前記予め設定された閾値以上であることに応答して、前記第１の検出結果と前記第２の検出結果に対して融合処理を行って、融合処理された結果を前記顔ビデオ内の顔の生体検出結果とするステップと、を含む、
ことを特徴とする請求項１に記載の方法。
【請求項５】
前記第１の検出結果と前記第２の検出結果に対して融合処理を行って、融合処理された結果を前記顔ビデオ内の顔の生体検出結果とするステップは、
前記第１の検出結果と前記第２の検出結果に対して平均計算を行い、前記平均計算によって得られた結果を前記顔ビデオ内の顔の生体検出結果とするステップ、または、
前記第１の検出結果と前記第２の検出結果に対して加重平均計算を行い、前記加重平均計算によって得られた結果を前記顔ビデオ内の顔の生体検出結果とするステップを含む、
ことを特徴とする請求項４に記載の方法。
【請求項６】
前記第１の検出結果と前記第２の検出結果は、それぞれ２項分類結果であり、前記２項分類結果の第１の分類結果が、前記顔ビデオ内の顔が生体である確率を表し、前記２項分類結果の第２の分類結果が、前記顔ビデオ内の顔が非生体である確率を表し、
前記第１の検出結果と前記第２の検出結果とに基づいて、前記顔ビデオ内の顔の生体検出結果を決定するステップは、
前記第２の検出結果の前記第１の分類結果が予め設定された閾値より小さいか否かを判断するステップと、
前記第２の検出結果の前記第１の分類結果が前記予め設定された閾値より小さいことに応答して、前記第２の検出結果に基づいて前記顔ビデオ内の顔の生体検出結果を決定するステップと、
前記第２の検出結果の前記第１の分類結果が前記予め設定された閾値以上であることに応答して、前記第１の検出結果と前記第２の検出結果に対して融合処理を行い、融合処理された結果を前記顔ビデオ内の顔の生体検出結果とするステップと、を含む、
ことを特徴とする請求項４に記載の方法。
【請求項７】
前記Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得するステップは、

前記Ｎフレームの顔画像に対してピクセル正規化処理を行って、Ｎフレームの正規化された顔画像を取得するステップと、
前記Ｎフレームの正規化された顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得するステップと、を含む、
ことを特徴とする請求項１に記載の方法。
【請求項８】
顔生体検出装置であって、
顔ビデオにおける連続したＮフレームの画像を取得し、前記Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得する取得モジュールであって、Ｎが１より大きい整数である取得モジュールと、
前記Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する第１の差分モジュールと、
前記Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得する第２の差分モジュールと、
前記Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、前記第２の差分画像に基づいて対応する第２の検出結果を決定する第１の決定モジュールと、
前記第１の検出結果と前記第２の検出結果とに基づいて、前記顔ビデオ内の顔の生体検出結果を決定する第２の決定モジュールと、を含む、
ことを特徴とする顔生体検出装置。
【請求項９】
前記第１の決定モジュールが、
前記Ｎ－１フレームの第１の差分画像に対して時空間特徴抽出を行って、前記Ｎ－１フレームの第１の差分画像の時空間特徴ベクトルを取得し、
前記時空間特徴ベクトルに対して加重計算を行って、対応する第１の加重結果を取得し、
前記第１の加重結果に対して正規化処理を行って、対応する第１の検出結果を取得する、
ことを特徴とする請求項８に記載の装置。
【請求項１０】
前記第１の決定モジュールが、
前記第２の差分画像に対して特徴抽出を行って、対応する第１の特徴ベクトルを取得し、
前記第１の特徴ベクトルに対して加重計算を行って、対応する第２の加重結果を取得し、
前記第２の加重結果に対して正規化処理を行って、対応する第２の検出結果を取得する、
ことを特徴とする請求項８に記載の装置。
【請求項１１】
前記第２の決定モジュールが、
前記第２の検出結果が予め設定された閾値より小さいか否かを判断する判断ユニットと、
前記第２の検出結果が前記予め設定された閾値より小さい場合、前記第２の検出結果に基づいて前記顔ビデオ内の顔の生体検出結果を決定する第１の決定ユニットと、
前記第２の検出結果が前記予め設定された閾値以上である場合、前記第１の検出結果と前記第２の検出結果に対して融合処理を行い、融合処理された結果を前記顔ビデオ内の顔の生体検出結果とする第２の決定ユニットと、を含む、
ことを特徴とする請求項８に記載の装置。
【請求項１２】
前記第２の決定ユニットが、
前記第１の検出結果と前記第２の検出結果に対して平均計算を行い、前記平均計算によって得られた結果を前記顔ビデオ内の顔の生体検出結果とし、または、
前記第１の検出結果と前記第２の検出結果に対して加重平均計算を行い、前記加重平均計算によって得られた結果を前記顔ビデオ内の顔の生体検出結果とする、
ことを特徴とする請求項１１に記載の装置。
【請求項１３】
前記第１の検出結果と前記第２の検出結果は、それぞれ２項分類結果であり、前記２項

分類結果の第１の分類結果が、前記顔ビデオ内の顔が生体である確率を表し、前記２項分類結果の第２の分類結果が、前記顔ビデオ内の顔が非生体である確率を表し、
前記判断ユニットが、
前記第２の検出結果の前記第１の分類結果が予め設定された閾値より小さいか否かを判断し、
前記第１の決定ユニットが、
前記第２の検出結果の前記第１の分類結果が前記予め設定された閾値より小さいことに応答して、前記第２の検出結果に基づいて前記顔ビデオ内の顔の生体検出結果を決定し、
前記第２の決定ユニットが、
前記第２の検出結果の前記第１の分類結果が前記予め設定された閾値以上であることに応答して、前記第１の検出結果と前記第２の検出結果に対して融合処理を行い、融合処理された結果を前記顔ビデオ内の顔の生体検出結果とする、
ことを特徴とする請求項１１に記載の装置。
【請求項１４】
前記装置は、
前記Ｎフレームの顔画像に対してピクセル正規化処理を行って、Ｎフレームの正規化された顔画像を取得する正規化モジュールをさらに含み、
前記第１の差分モジュールが、
前記Ｎフレームの正規化された顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する、
ことを特徴とする請求項１１に記載の装置。
【請求項１５】
電子機器であって、
少なくとも１つのプロセッサと、
前記少なくとも１つのプロセッサと通信可能に接続されるメモリと、を含み、
前記メモリには、前記少なくとも１つのプロセッサによって実行可能な命令が記憶されており、前記命令は、前記少なくとも１つのプロセッサが請求項１～７のいずれかに記載の方法を実行できるように、前記少なくとも１つのプロセッサによって実行される、
ことを特徴とする電子機器。
【請求項１６】
コンピュータ命令が記憶されている非一時的なコンピュータ読み取り可能な記憶媒体であって、
前記コンピュータ命令は、コンピュータに請求項１～７のいずれかに記載の方法を実行させる、
ことを特徴とする非一時的なコンピュータ読み取り可能な記憶媒体。
【請求項１７】
コンピュータプログラムであって、
前記コンピュータプログラムがプロセッサによって実行される場合、請求項１～７のいずれかに記載の方法が実現される、
ことを特徴とするコンピュータプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本開示は、人工知能技術の分野に関し、具体的には、コンピュータビジョンと深層学習技術の分野に関し、顔認識などのシーンに適用可能であり、特に顔生体検出方法、装置、電子機器及び記憶媒体に関する。
【背景技術】
【０００２】
顔生体検出は、画像が偽造された顔であるか否かを区別することであり、顔認識システムの基礎構成モジュールであり、顔認識システムの安全性を確保する。深層学習技術を用

いた顔ディープフェイクアルゴリズムは現在当分野の主流の方法である。しかしながら、いくつかのアプリケーションシーンでは、深層学習に基づく顔ディープフェイクアルゴリズムは汎化性が悪く、未知の偽造サンプルや方式に対する効果が低下するという問題があり、実用性能に影響を及ぼす。

【発明の概要】

【発明が解決しようとする課題】

【０００３】

本開示は、顔生体検出のための方法、装置、機器及び記憶媒体を提供する。

【課題を解決するための手段】

【０００４】

本開示の第１の態様によれば、

顔ビデオにおける連続したＮフレームの画像を取得し、前記Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得するステップであって、Ｎが１より大きい整数であるステップと、

前記Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得するステップと、

前記Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得するステップと、

前記Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、前記第２の差分画像に基づいて対応する第２の検出結果を決定するステップと、

前記第１の検出結果と前記第２の検出結果とに基づいて、前記顔ビデオ内の顔の生体検出結果を決定するステップと、を含む顔生体検出方法を提供する。

【０００５】

本開示の第２の態様によれば、

顔ビデオにおける連続したＮフレームの画像を取得し、前記Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得する取得モジュールであって、Ｎが１より大きい整数である取得モジュールと、

前記Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する第１の差分モジュールと、

前記Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得する第２の差分モジュールと、

前記Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、前記第２の差分画像に基づいて対応する第２の検出結果を決定する第１の決定モジュールと、

前記第１の検出結果と前記第２の検出結果とに基づいて、前記顔ビデオ内の顔の生体検出結果を決定する第２の決定モジュールと、を含む顔生体検出装置を提供する。

【０００６】

本開示の第３の態様によれば、電子機器を提供し、

少なくとも１つのプロセッサと、

前記少なくとも１つのプロセッサと通信可能に接続されるメモリと、を含み、

前記メモリには、前記少なくとも１つのプロセッサによって実行可能な命令が記憶されており、前記命令は、前記少なくとも１つのプロセッサが上記第１の態様に記載の方法を実行できるように、前記少なくとも１つのプロセッサによって実行される。

【０００７】

本開示の第４の態様によれば、コンピュータ命令が記憶されている非一時的なコンピュータ読み取り可能な記憶媒体を提供し、前記コンピュータ命令は、コンピュータに上記第１の態様に記載の方法を実行させる。

【０００８】

本開示の第５の態様によれば、コンピュータプログラムを提供し、前記コンピュータプログラムがプロセッサによって実行される場合、上記第１の態様に記載の方法が実現される。

【発明の効果】
【０００９】
本開示の技術案によれば、既存の顔偽造技術に対して、顔ビデオにおける連続した複数フレームの顔画像に対して前後フレームの差分を計算し、得られた差分画像に基づいて対応する検出結果を決定することにより、時間次元における顔画像の連続性に基づいて、対応する検出結果を決定することに相当し、時間次元における連続性が悪い顔偽造を効果的に防御することができる。また、顔ビデオにおける１フレーム目の画像と最後のフレームの画像との差分を計算して得られた差分画像も生体検出の根拠とし、顔ビデオにおける背景の変化の度合いも考慮されることに相当し、これにより、顔領域のみを処理して背景が変わらない顔偽造方式を効果的に防御することができ、さらに顔生体検出技術の性能を向上させ、顔認識システムの安全を確保するとともに、顔生体検出技術を基礎とする多くのアプリケーションの効果と体験を向上させることができる。
【００１０】
なお、この部分に記載の内容は、本開示の実施例の肝心または重要な特徴を特定することを意図しておらず、本開示の範囲を限定することも意図していないことを理解されたい。本開示の他の特徴は、以下の説明を通して容易に理解される。
【図面の簡単な説明】
【００１１】
図面は、本技術案をよりよく理解するために使用され、本開示を限定するものではない。
【図１】本開示の実施例によって提案される顔生体検出方法のフローチャートである。
【図２】本開示の実施例におけるビデオ内の顔の生体検出結果を決定するフローチャートである。
【図３】本開示の実施例によって提案される別の顔生体検出方法のフローチャートである。
【図４】本開示の実施例の顔生体検出モデルのネットワーク構造の概略図である。
【図５】本開示の実施例によって提案される顔生体検出装置の構成ブロック図である。
【図６】本開示の実施例によって提案される別の顔生体検出装置の構成ブロック図である。
【図７】本開示の実施例によって提案される電子機器の構成ブロック図である。
【発明を実施するための形態】
【００１２】
以下、図面と併せて本開示の例示的な実施例を説明し、理解を容易にするためにその中には本開示の実施例の様々な詳細事項が含まれており、それらは単なる例示的なものと見なされるべきである。したがって、当業者は、本開示の範囲及び精神から逸脱することなく、ここで説明される実施例に対して様々な変更と修正を行うことができることを認識されたい。同様に、わかりやすく簡潔にするために、以下の説明では、周知の機能及び構造の説明を省略する。
【００１３】
なお、顔生体検出は、画像が偽造された顔であるか否かを区別することであり、顔認識システムの基礎構成モジュールであり、顔認識システムの安全性を確保する。現在の顔生体検出方法は主にニューラルネットワークの深層学習方法を用いている。深層学習を用いた顔生体検出の主な方法は、畳み込みニューラルネットワークに基づく偽造ビデオ判別、長短期記憶人工ニューラルネットワークに基づく偽造ビデオ判別などの方法である。このような方法はニューラルネットワークを用いて顔の特徴抽出と分類を行う。深層学習方法は安定性の高い顔特徴を抽出することができ、より良い性能を有する。
【００１４】
しかしながら、従来技術では、顔生体検出の方法は、未知の偽造方式やサンプルに対する汎化性に限界がある。また、ビデオフレームのタイミング情報を考慮していないため、実際のシーンのビデオ偽造に対してロバスト性が悪く、認識効果が望ましくない。
【００１５】
また、本開示の発明者は、現在の顔偽造技術の研究により、現在主流の顔偽造技術であるＦａｃｅ２Ｆａｃｅ、ＤｅｅｐＦａｋｅｓ、ＦａｃｅＳｗａｐなどの多くは、シングル

フレーム処理に基づく方法であると発見され、つまり、他の人の顔を元の顔に貼り合わせて偽造を完了させ、生成された偽造ビデオが時間次元において連続性がなく、特に顔の動き中にあるフレームに偽造効果が悪い場合があり、実の顔は２フレーム間で大きく変化しない。また、ＦＯＭなどのような顔の表情のみを駆動する深層偽造技術は顔領域のみを処理し、背景を処理しない。

【００１６】

そのため、本開示は、上記の課題及び発見に基づいて、顔ビデオフレーム間の差分画像に基づいて生体検出結果を決定する顔生体検出方法、装置、電子機器及び記憶媒体を提案する。

【００１７】

本開示の実施例を容易に理解するために、まず、顔生体検出方法について説明する。

【００１８】

なお、本開示の技術案では、関係するユーザ個人情報の取得、保存、応用などは、いずれも関連する法律法規の規定に適合し、公序良俗に違反することはない。また、関係するユーザ個人情報は、ユーザの同意を得て取得、保存、適用されたものである。

【００１９】

図１は、本開示の実施例によって提案される顔生体検出方法のフローチャートである。なお、本開示の実施例の顔生体検出方法は、本開示の実施例の顔生体検出装置に適用することができ、当該装置は電子機器に配置されることができる。図１に示すように、当該方法は以下のステップ１０１～１０５を含む。

【００２０】

ステップ１０１において、顔ビデオにおける連続したＮフレームの画像を取得し、Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得し、Ｎが１より大きい整数である。

【００２１】

通常、顔認識を行う端末機器には、カメラが配置されることができ、顔認識を行う場合、端末機器のカメラをトリガーして撮影し、撮影されたビデオに基づいて顔認識を行う。本開示の実施例では、顔ビデオは、顔認識システムによって、端末に配置されたカメラを介して撮影されてもよい。なお、当該顔ビデオには顔情報が含まれる必要があり、一例として、端末カメラによって撮影されたビデオを検証処理を行うことができ、ビデオに顔情報が含まれていないことに応答して、再撮影して検証する必要がある。

【００２２】

また、顔ビデオのフレーム数がＮ以上である場合、その中の連続したＮフレームを顔ビデオにおける連続したＮフレームの画像として取得することができる。顔ビデオのフレーム数がＮである場合、当該ビデオに対応するＮフレームは、顔ビデオにおける連続したＮフレームの画像である。なお、当該Ｎフレームの画像はいずれも顔を含む画像である。

【００２３】

本開示の実施例では、後続の顔位置合わせ操作を容易にするために、Ｎフレームの画像を前処理して、顔に対応する領域を取得することができる。一例として、Ｎフレームの画像のうちの各枚の画像を検出モデルによって顔を検出して、各枚の画像内の顔の位置領域を取得することができる。ここで、検出モデルは、既存の顔検出モデルであり、顔の位置を検出することができる。

【００２４】

次に、検出された顔の位置に基づいて、Ｎフレームの画像に対して顔位置合わせを行って、Ｎフレームの顔画像を取得する。一例として、検出された顔領域に基づいて、顔キーポイント検出モデルによって顔キーポイントを検出して、顔キーポイントの座標値を取得することができる。ここで、顔検出ポイントモデルは既存のモデルである。顔区域が検出されたＮフレームの画像を当該モデルに入力して、それぞれが（ｘ１，ｙ１）…（ｘ７２，ｙ７２）である７２個の顔キーポイント座標を取得することができる。顔キーポイント座標値に基づいてターゲット顔を位置合わせして対応する顔画像を取得する。具体的には

、のの位置合わせの過程は、以下の通りであってもよい。上記７２個の顔キーポイント座標に基づいて横座標の最小値ｘｍｉｎ、横座標の最大値ｘｍａｘ、縦座標の最小値ｙｍｉｎ及び縦座標の最大値ｙｍａｘを取得し、当該４つの点の座標に基づいて１つの顔枠を決定し、決定された顔枠を２倍拡大してから顔画像を切り取り、サイズを２２４＊２２４に調整して、Ｎフレームの顔画像を取得することができる。

【００２５】

ステップ１０２において、Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する。

【００２６】

なお、実際の顔ビデオは、２フレーム間での変化が小さく、偽造された顔ビデオがシングルフレームのみに基づいて処理されるため、生成された偽造ビデオは時間次元における連続性が悪い。したがって、Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算し、すなわち、Ｎフレームの顔画像のうち、隣接する２フレーム間の差異を計算し、得られた差分画像は、時間次元における顔ビデオの連続性を表すことができ、さらに、当該ビデオ内の顔の真偽を判断することができる。

【００２７】

一例として、Ｎフレームの顔画像のうちの隣接する２フレームごとに後フレームと前フレームとの差を計算することができ、２フレームの顔画像ごとに差を計算して１枚の差分画像を取得することができるため、当該Ｎフレームの顔画像でＮ－１フレームの第１の差分画像を取得することができる。

【００２８】

ステップ１０３において、Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得する。

【００２９】

既存の顔偽造技術では、顔の表情のみを駆動する深層偽造技術があり、それは、顔領域のみを偽造し、背景情報が変わらないことに相当する。つまり、Ｎフレームの画像における背景情報が一致するか否かに基づいて、ビデオ内の顔の真偽を判断する根拠とすることができる。

【００３０】

本開示の実施例では、Ｎフレームの画像のうちのＮフレーム目の画像と１フレーム目の画像とを差分処理することによって第２の差分画像を取得し、すなわち、Ｎフレームの画像のうちのＮフレーム目の画像と１フレーム目の画像との差分を計算することによって第２の差分画像を取得する。ここで、第２の差分画像は、Ｎフレームの画像における背景情報の変化の度合いを表す。

【００３１】

ステップ１０４において、Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、第２の差分画像に基づいて対応する第２の検出結果を決定する。

【００３２】

なお、Ｎ－１フレームの第１の差分画像は、顔ビデオにおける顔情報の隣接する２フレーム間の変化を表すことができ、すなわち、隣接する２フレーム間に第１の差分画像で表される顔情報の変化が大きいほど、当該ビデオ内の顔が偽造される可能性が高くなる。第２の差分画像は、顔ビデオにおける背景情報の変化を表すことができ、すなわち、第２の差分画像で表される顔ビデオにおける背景情報の変化が大きいほど、当該ビデオ内の顔が偽造される可能性が高くなる。

【００３３】

一例として、Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定する実現形態は以下の通りであってもよい。Ｎ－１フレームの第１の差分画像に対して時空間特徴抽出を行って、Ｎ－１フレームの第１の差分画像の時空間特徴ベクトルを取得する。例えば、トレーニングされた３Ｄ畳み込みニューラルネットワークによってＮ－１フレームの第１差分画像に対して時空間特徴抽出を行うことができる。取得された時空間

特徴ベクトルに対して加重計算を行って、時空間特徴情報を統合して、対応する第１の加重結果を取得する。得られた第１の加重結果に対して正規化処理を行って、対応する第１の検出結果を取得する。また、第２の差分画像に基づいて対応する第２の検出結果を決定する実現形態は以下の通りであってもよい。第２の差分画像に対して特徴抽出を行って、対応する第１の特徴ベクトルを取得する。例えば、畳み込みニューラルネットワークによって第２の差分画像に対して特徴抽出を行う。第１の特徴ベクトルに対して加重計算を行って、対応する第２の加重結果を取得する。第２の加重結果に対して正規化処理を行って、対応する第２の検出結果を取得する。ここで、第１の検出結果と第２の検出結果は、当該ビデオ内の顔が生体である確率であってもよいし、当該ビデオ内の顔が生体である確率と当該ビデオ内の顔が生体ではない確率を表す２項分類結果であってもよいし、または検出結果を表すことができる他の形態であってもよい。

【００３４】

別の例として、第１の差分画像をスコア分割することができ、スコアが高いほど隣接する２フレームの画像における顔情報の変化が大きくなることを示し、Ｎ－１フレームの第１の差分画像のスコアに基づいて当該ビデオ内の顔が生体である確率を計算し、当該確率値を第１の検出結果とする。同時に、第２の差分画像もスコア分割して、スコアが高いほど、背景情報の変化が大きくなることを示し、当該スコアを当該ビデオ内の顔が生体である確率に対応させて、第２の検出結果を取得する。

【００３５】

ステップ１０５において、第１の検出結果と第２の検出結果とに基づいて、ビデオ内の顔の生体検出結果を決定する。

【００３６】

なお、現在の顔偽造技術には、シングルフレームの画像に対する顔貼り合わせによって顔偽造を行う方式も、顔領域のみを偽造して背景が出ない方式もあるため、第１の検出結果と第２の検出結果とに基づいてビデオ内の顔の生体検出結果を決定することにより、顔生体検出の効果と精度を効果的に向上させることができる。

【００３７】

一例として、第１の検出結果と第２の検出結果とを融合処理し、例えば平均化処理、加重平均化処理などを行い、融合処理されたされた結果をビデオ内の顔の生体検出結果とすることができる。

【００３８】

本開示の実施例によって提案される顔生体検出方法によれば、既存の顔偽造技術に対して、顔ビデオにおける連続した複数フレームの顔画像に対して前後フレームの差分を計算することにより、得られた差分画像に基づいて対応する検出結果を決定することにより、時間次元における顔画像の連続性に基づいて、対応する検出結果を決定することに相当し、時間次元における連続性が悪い顔偽造を効果的に防御することができる。また、顔ビデオにおける１フレーム目の画像と最後のフレームの画像との差分を計算して得られた差分画像も生体検出の根拠とし、顔ビデオにおける背景の変化の度合いも考慮されることに相当し、これにより、顔領域のみを処理して背景が変わらない顔偽造方式を効果的に防御することができ、さらに顔生体検出技術の性能を向上させ、顔認識システムの安全を確保するとともに、顔生体検出技術を基礎とする多くのアプリケーションの効果と体験を向上させることができる。

【００３９】

ビデオ内の顔の生体検出結果を効果的に決定するために、上記の実施例に加えて、本開示は、別の実施例を提案する。

【００４０】

図２は、本開示の実施例におけるビデオ内の顔の生体検出結果を決定するフローチャートである。図２に示すように、ビデオ内の顔の生体検出結果を決定する実現形態は以下のステップ２０１～２０３の通りであってもよい。

【００４１】

ステップ２０１において、第２の検出結果が予め設定された閾値より小さいか否かを判断する。

【００４２】

なお、第２の検出結果は第２の差分画像に基づいて得られたものであり、すなわち、第２の差分画像を計算することにより、連続したＮフレームの画像における背景情報の変化状況を取得し、背景情報の変化状況に基づいて第２の検出結果を取得する。第２の検出結果において、当該ビデオ内の顔が生体である確率が低い、すなわち、当該ビデオにおける背景情報がほとんど変わらない場合、現在のビデオ内の顔が偽造されたものであると決定することができる。

【００４３】

本開示の実施例では、第２の検出結果は、当該ビデオ内の顔が生体である確率であり、予め設定された閾値とは、第２の検出結果における当該ビデオ内の顔が生体である確率の閾値を指し、この値の大きさは、実際の状況に応じて複数回の実験によって決定することができる。なお、当該閾値は、第２の検出結果に基づいて、生体検出結果を直接決定できるか否かの限界とすることができ、第２の検出結果が当該閾値より小さいことに応答して、当該ビデオ内の顔が生体である確率が極めて小さいことを示し、当該ビデオ内の顔の生体検出結果が非生体であると直接決定することができる。第２の検出結果が当該閾値以上であることに応答して、当該ビデオ内の顔が生体、または、偽造されたものである可能性があることを示し、第１の検出結果と組み合わせてさらに決定する必要がある。このように、第２の検出結果を顔の生体検出結果を決定する初歩的な判断とすることで、ビデオ内の収束した生体検出結果を決定する計算量を低減させ、結果出力効率を向上させることができる。

【００４４】

ステップ２０２において、第２の検出結果が予め設定された閾値より小さいことに応答して、第２の検出結果に基づいてビデオ内の顔の生体検出結果を決定する。

【００４５】

つまり、第２の検出結果に対応する当該ビデオ内の顔が生体である確率が予め設定された閾値より小さいことに応答して、第２の検出結果に基づいてビデオ内の顔の生体検出結果を直接決定することができる。

【００４６】

例えば、得られた第２の検出結果が、当該ビデオ内の顔が生体である確率が０．０５であり、予め設定された閾値が０．１であり、すなわち第２の検出結果が予め設定された閾値より小さい場合、第２の検出結果に基づいて、ビデオ内の顔の生体検出結果が非生体であると直接決定することができる。

【００４７】

ステップ２０３において、第２の検出結果が予め設定された閾値以上であることに応答して、第１の検出結果と第２の検出結果を融合処理し、融合処理された結果をビデオ内の顔の生体検出結果とする。

【００４８】

つまり、第２の検出結果に対応する当該ビデオ内の顔が生体である確率が予め設定された閾値以上であることに応答して、この場合、ビデオ内の顔が生体または偽造されたものである可能性があり、すなわち、第２の検出結果のみに基づいてビデオ内の顔の生体検出結果を決定することができず、第１の検出結果と組み合わせてさらに決定する必要がある。

【００４９】

本開示の実施例では、第１の検出結果と第２の検出結果を融合処理し、融合処理された結果をビデオ内の顔の生体検出結果とする実現形態は以下の通りであってもよい。第１の検出結果と第２の検出結果に対して平均計算を行い、平均計算によって得られた結果をビデオ内の顔の生体検出結果とする。または、第１の検出結果と第２の検出結果に対して加重平均計算を行い、加重平均計算によって得られた結果をビデオ内の顔の生体検出結果とする。

【００５０】
別の実施形態として、第１の検出結果と第２の検出結果は２項分類結果であってもよい。当該２項分類結果の第１の分類結果はビデオ内の顔が生体である確率を表し、当該２項分類結果の第２の分類結果はビデオ内の顔が非生体である確率を表す。上記ステップ２０１は、第２の検出結果の第１の分類結果が予め設定された閾値より小さいか否かを判断する。ステップ２０２は、第２の検出結果の第１の分類結果が予め設定された閾値より小さいことに応答して、第２の検出結果に基づいてビデオ内の顔の生体検出結果を決定する。ステップ２０３は、第２の検出結果の第１の分類結果が予め設定された閾値以上であることに応答して、第１の検出結果と第２の検出結果を融合処理し、融合処理された結果をビデオ内の顔の生体検出結果とする。

【００５１】
本開示の実施例の顔生体検出方法によれば、ビデオ内の顔の生体検出結果を決定する場合、まず、第２の検出結果を予め設定された閾値と比較し、ビデオ内の顔の生体検出結果を決定する初歩的な判断とし、第２の検出結果が予め設定された閾値より小さいことに応答して、直接第２の検出結果に基づいて生体検出結果として、すなわち、当該ビデオ内の顔生体検出結果が非生体であると直接判断し、そうでなければ、第１の検出結果と組み合わせてさらに決定する。このように、ビデオ内の収束した生体検出結果を決定する計算量を効果的に低減させ、結果出力効率を向上させることができる。

【００５２】
画像ピクセル値に対する光などの干渉を低減させ、幾何学的変換等による画像への影響を除去するために、本開示の実施例では、画像の正規化処理が追加されている。

【００５３】
図３は、本開示の実施例によって提案される別の顔生体検出方法のフローチャートである。上記の実施例に加えて、図３に示すように、当該方法は以下のステップ３０６をさらに含む。

【００５４】
ステップ３０６において、Ｎフレームの顔画像に対してピクセル正規化処理を行って、Ｎフレームの正規化された顔画像を取得する。

【００５５】
なお、正規化の基本思想は、画像の不変モーメントを用いて、他の変換関数による画像変換への影響を除去できる１組のパラメータを探し、すなわち、アフィン変換に抵抗するために一意の標準形式に変換することである。画像のピクセル正規化処理は、画像を幾何学的変換の攻撃に抵抗させ、光の不均一によるピクセル値の干渉を低減させることができる。

【００５６】
一例として、Ｎフレームの顔画像に対してピクセル正規化処理を行う実現形態は以下の通りであってもよい。画像の各ピクセルのピクセル値から１２８を引いて、それぞれ２５６で除算し、各ピクセルのピクセル値を［－０．５、０．５］の間にする。

【００５７】
ステップ３０２において、Ｎフレームの正規化された顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する。

【００５８】
なお、図３のステップ３０１～３０５は、図１のステップ１０１～１０５と同じ実現形態を有し、ここでは説明を省略する。

【００５９】
本開示の実施例の顔生体検出方法によれば、顔画像にピクセル正規化処理を追加することにより、光の不均一によるピクセルの干渉を効果的に低減することができ、画像が幾何学的変換の攻撃に抵抗でき、顔生体検出の精度を向上させることができる。

【００６０】
上記の実施例に基づいて、Ｎ－１フレームの第１の差分画像に基づいて対応する第１の

検出結果を決定し、第２の差分画像に基づいて対応する第２の検出結果を決定するプロセスは、顔生体検出モデルによって実現することができる。

【００６１】

図４は当該顔生体検出モデルのネットワーク構造の概略図である。図４に示すように、顔生体検出モデルは、時空間特徴抽出層４０１、第１の完全接続層４０２、第１の正規化層４０３、画像特徴抽出層４０４、第２の完全接続層４０５及び第２の正規化層４０６を含む。当該モデルは、事前に、生体ビデオの連続したＮフレームの画像サンプルと、偽造された顔ビデオの連続したＮフレームの画像サンプルデータでトレーニングして得られる。

【００６２】

具体的には、Ｎ－１フレームの第１の差分画像を時空間特徴抽出層４０１に入力して時空間特徴抽出を行って、Ｎ－１フレームの第１の差分画像の時空間特徴ベクトルを取得する。ここで、時空間特徴抽出層は、Ｉ３Ｄ畳み込みニューラルネットワークであってもよいし、時空間特徴を抽出できる他のネットワークモデルであってもよい。取得された時空間特徴ベクトルを第１の完全接続層４０２に入力して加重計算を行って、対応する第１の加重結果を取得する。第１の加重結果を第１の正規化層４０３に入力して正規化処理を行って、対応する第１の検出結果を取得する。ここで、正規化層４０３はｓｏｆｔｍａｘ層であってもよい。

【００６３】

また、第２の差分画像を画像特徴抽出層４０４に入力して特徴抽出を行って、対応する第１の特徴ベクトルを取得する。ここで、画像特徴抽出層４０４はＭｏｂｉｌｅＮｅｔＶ２畳み込みニューラルネットワークであってもよいし、他の畳み込みニューラルネットワークモデルであってもよい。取得された第１の特徴ベクトルを第２の完全接続層４０５に入力して加重計算を行って、対応する第２の加重結果を取得する。取得された第２の加重結果を第２の正規化層４０６に入力して正規化処理を行って、対応する第２の検出結果を取得する。

【００６４】

本開示の実施例では、当該顔生体検出モデルのトレーニング方式は以下の通りであってもよい。顔ビデオの連続したＮフレームの画像サンプルを取得する。Ｎフレームの画像サンプルは生体顔ビデオのＮフレームの連続した画像サンプルと偽造された顔ビデオのＮフレームの連続した画像サンプルを含む。Ｎフレームの画像サンプルに対して顔位置合わせ処理を行って、Ｎフレームの顔画像サンプルを取得する。Ｎフレームの顔画像サンプルに対してピクセル正規化操作を行って、ランダムデータ拡張処理を行う。Ｎフレームの顔画像サンプルにおける各組のＮフレームの顔画像に対して、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像サンプルを取得する。Ｎフレームの画像サンプルにおける各組のＮフレームの画像に対して、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像サンプルを取得する。Ｎ－１フレームの第１の差分画像サンプルを顔生体検出モデルの時空間特徴抽出層に入力して時空間特徴抽出を行って、時空間特徴抽出層で抽出された情報を第１の完全接続層に入力して加重計算を行って、得られた加重計算結果に対して第１の正規化層によって正規化処理を行って、第１の予測結果を取得する。第２の差分画像サンプルを画像特徴抽出層に入力して特徴抽出を行って、抽出された情報を第２の完全接続層に入力して加重計算を行って、得られた加重計算結果に対して第２の正規化層によって正規化処理を行って、第２の予測結果を取得する。第１の予測結果と第２の予測結果とに基づいて対応する損失値を計算して、顔生体検出モデルをトレーニングする。

【００６５】

本開示の実施例によって提案される顔生体検出方法によれば、顔生体検出モデルによって、取得された差分画像に対してそれぞれ特徴抽出及び分類計算を行って、対応する検出結果を取得し、顔生体検出モデルを使用して生体ビデオと偽造されたビデオからより判別性のある特徴を抽出して分類検出を行うことに相当し、顔生体検出の汎化性と精度を効果的に向上させることができ、検出効率を向上させることもできる。また、当該顔検出モデ

ルを、生体ビデオの連続したフレームの画像サンプルと偽造された顔ビデオの連続したフレームの画像サンプルから得られた差分画像でトレーニングすることにより、ネットワークトレーニングの収束速度を速め、モデルトレーニングの効率を向上させることができる。

【００６６】

上記の実施例を実現するために、本開示は、顔生体検出装置を提案する。

【００６７】

図５は、本開示の実施例によって提案される顔生体検出装置の構成ブロック図である。図５に示すように、当該装置は、

顔ビデオにおける連続したＮフレームの画像を取得し、Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得する取得モジュール５１０であって、Ｎが１より大きい整数である取得モジュール５１０と、

Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する第１の差分モジュール５２０と、

Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得する第２の差分モジュール５３０と、

Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、第２の差分画像に基づいて対応する第２の検出結果を決定する第１の決定モジュール５４０と、

第１の検出結果と第２の検出結果とに基づいて、ビデオ内の顔の生体検出結果を決定する第２の決定モジュール５５０と、を含む。

【００６８】

本開示のいくつかの実施例では、第１の決定モジュール５４０は、具体的に、

Ｎ－１フレームの第１の差分画像に対して時空間特徴抽出を行って、Ｎ－１フレームの第１の差分画像の時空間特徴ベクトルを取得し、

時空間特徴ベクトルに対して加重計算を行って、対応する第１の加重結果を取得し、

第１の加重結果に対して正規化処理を行って、対応する第１の検出結果を取得する。

【００６９】

本開示のいくつかの実施例では、第１の決定モジュール５４０は、さらに、

第２の差分画像に対して特徴抽出を行って、対応する第１の特徴ベクトルを取得し、

第１の特徴ベクトルに対して加重計算を行って、対応する第２の加重結果を取得し、

第２の加重結果に対して正規化処理を行って、対応する第２の検出結果を取得する。

【００７０】

さらに、本開示のいくつかの実施例では、第２の決定モジュール５５０は、

第２の検出結果が予め設定された閾値より小さいか否かを判断する判断ユニット５５１と、

第２の検出結果が予め設定された閾値より小さい場合、第２の検出結果に基づいてビデオ内の顔の生体検出結果を決定する第１の決定ユニット５５２と、

第２の検出結果が予め設定された閾値以上である場合、第１の検出結果と第２の検出結果を融合処理し、融合処理された結果をビデオ内の顔の生体検出結果とする第２の決定ユニット５５３と、を含む。

【００７１】

ここで、第２の決定ユニット５５３は、具体的に、

第１の検出結果と第２の検出結果に対して平均計算を行い、平均計算によって得られた結果をビデオ内の顔の生体検出結果とし、または、

第１の検出結果と第２の検出結果に対して加重平均計算を行い、加重平均計算によって得られた結果をビデオ内の顔の生体検出結果とする。

【００７２】

本開示の別の実施例では、第１の検出結果と第２の検出結果は、それぞれ２項分類結果であり、２項分類結果の第１の分類結果がビデオ内の顔が生体である確率を表し、２項分類結果の第２の分類結果がビデオ内の顔が非生体である確率を表し、ここで、判断ユニット５５１は、具体的に、

第２の検出結果の第１の分類結果が予め設定された閾値より小さいか否かを判断し、
第１の決定ユニット５５２は、具体的に、
第２の検出結果の第１の分類結果が予め設定された閾値より小さいことに応答して、第２の検出結果に基づいてビデオ内の顔の生体検出結果を決定し、
第２の決定ユニット５５３は、具体的に、
第２の検出結果の第１の分類結果が予め設定された閾値以上であることに応答して、第１の検出結果と第２の検出結果を融合処理し、融合処理された結果をビデオ内の顔の生体検出結果とする。

【００７３】

本開示の実施例によって提案される顔生体検出装置によれば、既存の顔偽造技術に対して、顔ビデオにおける連続した複数フレームの顔画像に対して前後フレームの差分を計算し、得られた差分画像に基づいて対応する検出結果を決定することにより、時間次元における顔画像の連続性に基づいて、対応する検出結果を決定することに相当し、時間次元における連続性が悪い顔偽造を効果的に防御することができる。また、顔ビデオにおける１フレーム目の画像と最後のフレームの画像との差分を計算して得られた差分画像も生体検出の根拠とし、顔ビデオにおける背景の変化の度合いも考慮されることに相当し、これにより、顔領域のみを処理して背景が変わらない顔偽造方式を効果的に防御することができ、さらに顔生体検出技術の性能を向上させ、顔認識システムの安全を確保するとともに、顔生体検出技術を基礎とする多くのアプリケーションの効果と体験を向上させることができる。

【００７４】

光などによる画像ピクセル値への干渉を低減し、幾何学的変換等による画像への影響を除去するために、本開示は別の顔生体検出装置を提案する。

【００７５】

図６は、本開示の実施例によって提案される別の顔生体検出装置の構成ブロック図である。上記の実施例に加えて、図６に示すように、当該装置は、
Ｎフレームの顔画像に対してピクセル正規化処理を行って、Ｎフレームの正規化された顔画像を取得する正規化モジュール６６０をさらに含み、
ここで、第１の差分モジュール６２０は、具体的に、
Ｎフレームの正規化された顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する。

【００７６】

なお、図６の６１０～６５０は、図５の５１０～５５０と同じ機能および構造を有しており、ここでは説明を省略する。

【００７７】

上記実施例の装置について、各モジュールが操作を実行する具体的な方式は、当該方法に関する実施例においてすでに詳細に説明されているが、ここでは説明を省略する。

【００７８】

本開示の実施例の顔生体検出装置によれば、顔画像にピクセル正規化処理を追加することにより、光の不均一によるピクセルの干渉を効果的に低減することができ、画像が幾何学的変換の攻撃に抵抗でき、顔生体検出の精度を向上させることができる。

【００７９】

本開示の実施例によれば、本開示は、電子機器、読み取り可能な記憶媒体、及びコンピュータプログラムをさらに提供する。

【００８０】

図７に示すように、それは本開示の実施例の顔生体検出方法に係る電子機器のブロック図である。電子機器は、ラップトップコンピュータ、デスクトップコンピュータ、ワークステーション、パーソナルデジタルアシスタント、サーバ、ブレードサーバ、メインフレームコンピュータ、および他の適切なコンピュータなどの様々な形態のデジタルコンピュータを表すことを目的とする。電子機器は、パーソナルデジタル処理、携帯電話、スマー

トフォン、ウェアラブルデバイス、および他の同様のコンピューティングデバイスなどの様々な形態のモバイルデバイスを表すこともできる。本明細書で示される部品、それらの接続と関係、およびそれらの機能は、単なる例であり、本明細書の説明および／または求められる本開示の実現を制限することを意図したものではない。

【００８１】

図７示すように、当該電子機器は、１つ又は複数のプロセッサ７０１と、メモリ７０２と、高速インターフェースと低速インターフェースを含む、各コンポーネントを接続するためのインターフェースと、を含む。各コンポーネントは、異なるバスで相互に接続され、共通のマザーボードに取り付けられてもよいし、又は必要に応じて他の形態で取り付けられてもよい。プロセッサは、外部入力／出力装置（インターフェースに結合されたディスプレイデバイスなど）にＧＵＩの図形情報をディスプレイするためにメモリに記憶されている命令を含む、電子機器内に実行される命令を処理することができる。他の実施形態では、必要であれば、複数のプロセッサ及び／又は複数のバスを、複数のメモリと複数のメモリとともに使用することができる。同様に、複数の電子機器を接続することができ、各電子機器は、部分的な必要な操作（例えば、サーバアレイ、ブレードサーバ、又はマルチプロセッサシステムとする）を提供する。図７では、１つのプロセッサ７０１を例とする。

【００８２】

メモリ７０２は、本開示によって提供される非一時的なコンピュータ読み取り可能な記憶媒体である。前記メモリには、少なくとも１つのプロセッサが本開示によって提供される顔生体検出方法を実行するように、少なくとも１つのプロセッサによって実行される命令が記憶されている。本開示の非一時的なコンピュータ読み取り可能な記憶媒体には、本開示によって提供される顔生体検出方法をコンピュータに実行させるためのコンピュータ命令が記憶されている。

【００８３】

メモリ７０２は、非一時的なコンピュータ読み取り可能な記憶媒体として、本開示の実施例における顔生体検出方法に対応するプログラム命令／モジュール（例えば、図５に示す第１取得モジュール５１０、第１の差分モジュール５２０、第２の差分モジュール５３０、第１の決定モジュール５４０及び第２の決定モジュール５５０）のような、非一時的なソフトウェアプログラム、非一時的なコンピュータ実行可能なプログラム及びモジュールを記憶するために使用される。プロセッサ７０１は、メモリ７０２に記憶されている非一時的なソフトウェアプログラム、命令及びモジュールを実行することにより、サーバの様々な機能アプリケーション及びデータ処理を実行し、すなわち、上記の方法の実施例における顔生体検出方法を実現する。

【００８４】

本開示の実施例は、コンピュータプログラムを提供し、当該コンピュータプログラムがプロセッサによって実行される場合、上記の方法の実施例の顔生体検出方法が実現される。

【００８５】

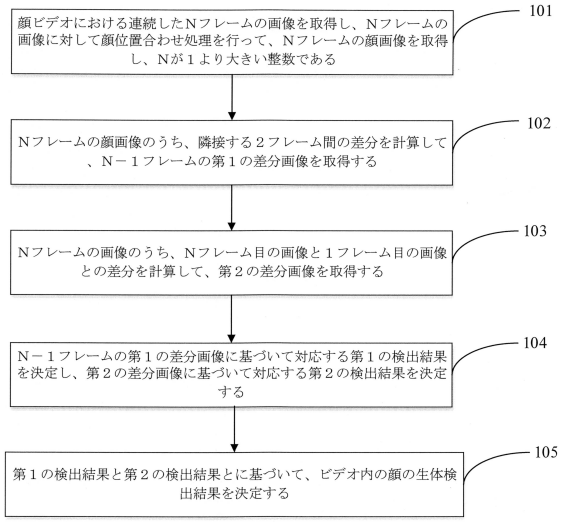
メモリ７０２は、記憶プログラム領域及び記憶データ領域を含むことができる。記憶プログラム領域は、オペレーティングシステム、少なくとも１つの機能に必要なアプリケーションを記憶することができる。記憶データ領域は、顔生体検出方法に係る電子機器の使用によって作成されたデータなどを記憶することができる。また、メモリ７０２は、高速ランダムアクセスメモリを含むことができ、少なくとも１つの磁気ディスクストレージデバイス、フラッシュメモリデバイス、又は他の非一時的なソリッドステートストレージデバイスなどの非一時的なメモリをさらに含むことができる。いくつかの実施例では、メモリ７０２は、プロセッサ７０１に対して遠隔に設定されたメモリを選択的に含むことができ、これらの遠隔メモリは、ネットワークを介して顔生体検出方法の電子機器に接続されることができる。上記ネットワークの例は、インターネット、イントラネット、ローカルエリアネットワーク、モバイル通信ネットワーク及びそれらの組み合わせを含むが、これらに限定されない。

【００８６】
顔生体検出方法の電子機器は、入力装置７０３と出力装置７０４とをさらに含むことができる。プロセッサ７０１、メモリ７０２、入力装置７０３、及び出力装置７０４は、バスまたは他の方式で接続することができ、図７では、バスを介して接続することを例に挙げる。
【００８７】
入力装置７０３は、入力された数字又は文字情報を受信し、顔生体検出方法の電子機器のユーザ設定及び機能制御に関するキー信号入力を生成することができ、例えば、タッチスクリーン、キーパッド、マウス、トラックパッド、タッチパッド、指示棒、１つ又は複数のマウスボタン、トラックボール、ジョイスティックなどの入力装置である。出力装置７０４は、ディスプレイデバイス、補助照明デバイス（例えば、ＬＥＤ）、及び触覚フィードバックデバイス（例えば、振動モータ）などを含むことができる。当該ディスプレイデバイスは、液晶ディスプレイ（ＬＣＤ）、発光ダイオード（ＬＥＤ）ディスプレイ、及びプラズマディスプレイを含むことができるが、これらに限定されない。いくつかの実施形態では、ディスプレイデバイスは、タッチスクリーンであってもよい。
【００８８】
本明細書で説明されるシステムと技術の様々な実施形態は、デジタル電子回路システム、集積回路システム、特定用途向けＡＳＩＣ（特定用途向け集積回路）、コンピュータハードウェア、ファームウェア、ソフトウェア、及び／又はそれらの組み合わせで実現することができる。これらの様々な実施形態は、１つ又は複数のコンピュータプログラムで実施されることを含むことができ、当該１つ又は複数のコンピュータプログラムは、少なくとも１つのプログラマブルプロセッサを含むプログラム可能なシステムで実行および／または解釈されることができ、当該プログラマブルプロセッサは、特定用途向け又は汎用プログラマブルプロセッサであってもよく、ストレージシステム、少なくとも１つの入力装置、および少なくとも１つの出力装置からデータおよび命令を受信し、データおよび命令を当該ストレージシステム、当該少なくとも１つの入力装置、および当該少なくとも１つの出力装置に伝送することができる。
【００８９】
これらのコンピューティングプログラム（プログラム、ソフトウェア、ソフトウェアアプリケーション、又はコードとも呼ばれる）は、プログラマブルプロセッサの機械命令を含み、これらのコンピュータプログラムは、高度のプロセス及び／又は対象指向プログラミング言語、及び／又はアセンブリ／機械言語で実施される。本明細書で使用されるような、「機械読み取り可能な媒体」及び「コンピュータ読み取り可能な媒体」という用語は、機械命令及び／又はデータをプログラマブルプロセッサに提供するために使用される任意のコンピュータプログラム製品、機器、及び／又は装置（例えば、磁気ディスク、光ディスク、メモリ、プログラマブルロジックデバイス（ＰＬＤ））を指し、機械読み取り可能な信号である機械命令を受信する機械読み取り可能な媒体を含む。「機械読み取り可能な信号」という用語は、機械命令及び／又はデータをプログラマブルプロセッサに提供するための任意の信号を指す。
【００９０】
ユーザとのインタラクションを提供するために、ここで説明されるシステム及び技術をコンピュータ上で実施することができ、当該コンピュータは、ユーザに情報を表示するためのディスプレイ装置（例えば、ＣＲＴ（陰極線管）又はＬＣＤ（液晶ディスプレイ）モニタ）と、キーボード及びポインティングデバイス（例えば、マウス又はトラックボール）とを有し、ユーザは、当該キーボード及び当該ポインティングデバイスによって入力をコンピュータに提供することができる。他の種類の装置も、ユーザとのインタラクションを提供することができ、例えば、ユーザに提供されるフィードバックは、任意の形態の感覚フィードバック（例えば、視覚フィードバック、聴覚フィードバック、又は触覚フィードバック）であってもよく、任意の形式（音響入力と、音声入力、または、触覚入力とを含む）でユーザからの入力を受信することができる。

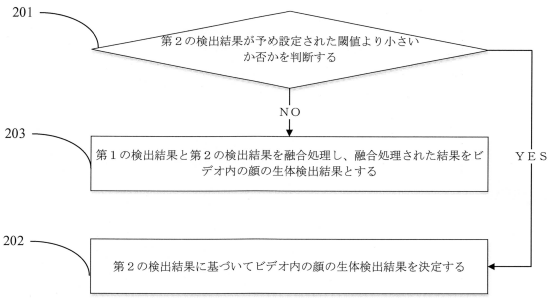
【００９１】
ここで説明されるシステムおよび技術は、バックエンドコンポーネントを含むコンピューティングシステム（例えば、データサーバとする）、又はミドルウェアコンポーネントを含むコンピューティングシステム（例えば、アプリケーションサーバ）、又はフロントエンドコンポーネントを含むコンピューティングシステム（例えば、グラフィカルユーザインタフェース又はウェブブラウザを有するユーザコンピュータ、ユーザは、当該グラフィカルユーザインタフェース又は当該ウェブブラウザによってここで説明されるシステムおよび技術の実施形態とインタラクションできる）、又はこのようなバックエンドコンポーネントと、ミドルウェアコンポーネントと、フロントエンドコンポーネントのいずれかの組み合わせを含むコンピューティングシステムで実行することができる。任意の形態又は媒体のデジタルデータ通信（例えば、通信ネットワーク）によってシステムのコンポーネントを相互に接続することができる。通信ネットワークとしての例は、ローカルエリアネットワーク（ＬＡＮ）と、ワイドエリアネットワーク（ＷＡＮ）と、インターネットと、ブロックチェーンネットワークを含む。
【００９２】
コンピュータシステムは、クライアントとサーバを含むことができる。クライアントとサーバは、一般に、互いに離れており、通常に通信ネットワークを介してインタラクションする。対応するコンピュータ上で実行され、互いにクライアント－サーバ関係を有するコンピュータプログラムによってクライアントとサーバとの関係が生成される。サーバはクラウドサーバであってもよく、クラウドコンピューティングサーバまたはクラウドホストとも呼ばれ、クラウドコンピューティングサービスシステムにおける１つのホスト製品であり、従来の物理ホストとＶＰＳサービス（「Ｖｉｒｔｕａｌ Ｐｒｉｖａｔｅ Ｓｅｒｖｅｒ」，または「ＶＰＳ」と省略する）に存在する管理の難しさ、ビジネス拡張性の弱いという欠陥を解決した。サーバは分散システムのサーバであってもよく、ブロックチェーンを組み合わせたサーバであってもよい。
【００９３】
なお、上記に示される様々な形式のフローを使用して、ステップを並べ替え、追加、又は削除することができることを理解されたい。例えば、本開示に記載の各ステップは、並列に実行されてもよいし、順次実行されてもよいし、異なる順序で実行されてもよいが、本開示で開示されている技術案が所望の結果を実現することができれば、本明細書では限定されない。
【００９４】
上記具体的な実施形態は、本開示の保護範囲を制限するものではない。当業者は、設計要求と他の要因に応じて、様々な修正、組み合わせ、サブコンビネーション、及び代替を行うことができると理解されたい。本開示の精神と原則内で行われる任意の修正、同等の置換、及び改善などは、いずれも本開示の保護範囲内に含まれなければならない。

【図面】

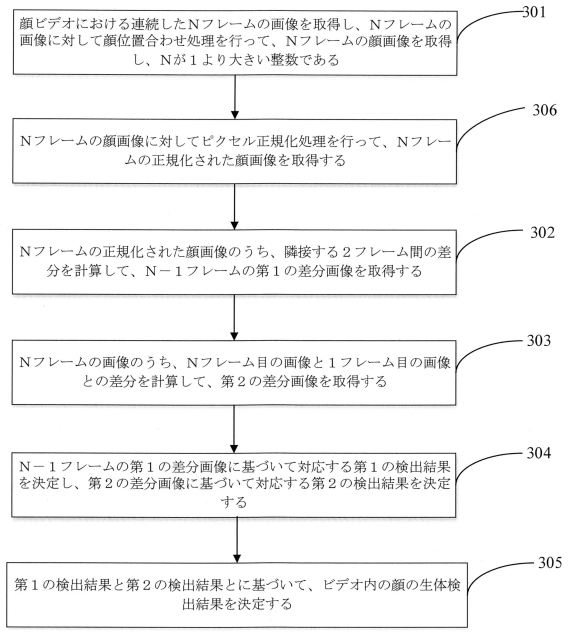
【図１】


顔ビデオにおける連続したＮフレームの画像を取得し、Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得し、Ｎが１より大きい整数である
101
Ｎフレームの顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する
102
Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得する
103
Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、第２の差分画像に基づいて対応する第２の検出結果を決定する
104
第１の検出結果と第２の検出結果とに基づいて、ビデオ内の顔の生体検出結果を決定する
105

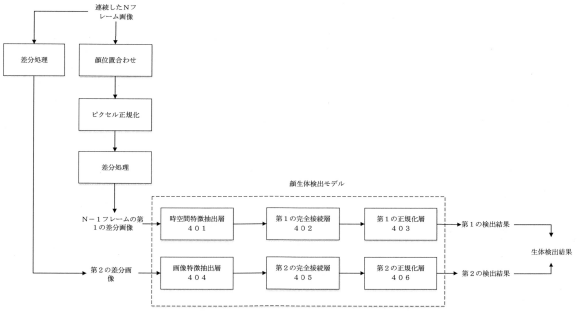
【図２】


201
第２の検出結果が予め設定された閾値より小さいか否かを判断する
NO
YES
203
第１の検出結果と第２の検出結果を融合処理し、融合処理された結果をビデオ内の顔の生体検出結果とする
202
第２の検出結果に基づいてビデオ内の顔の生体検出結果を決定する

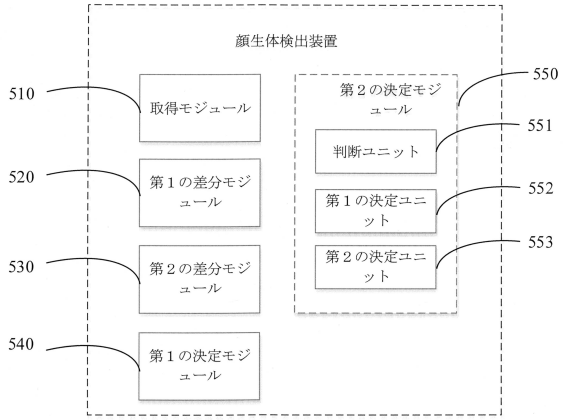
【図３】


顔ビデオにおける連続したＮフレームの画像を取得し、Ｎフレームの画像に対して顔位置合わせ処理を行って、Ｎフレームの顔画像を取得し、Ｎが１より大きい整数である
301
Ｎフレームの顔画像に対してピクセル正規化処理を行って、Ｎフレームの正規化された顔画像を取得する
306
Ｎフレームの正規化された顔画像のうち、隣接する２フレーム間の差分を計算して、Ｎ－１フレームの第１の差分画像を取得する
302
Ｎフレームの画像のうち、Ｎフレーム目の画像と１フレーム目の画像との差分を計算して、第２の差分画像を取得する
303
Ｎ－１フレームの第１の差分画像に基づいて対応する第１の検出結果を決定し、第２の差分画像に基づいて対応する第２の検出結果を決定する
304
第１の検出結果と第２の検出結果とに基づいて、ビデオ内の顔の生体検出結果を決定する
305

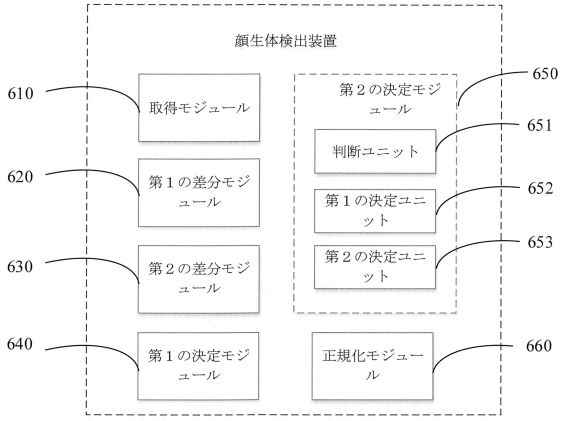
【図４】


連続したＮフレーム画像
差分処理
顔位置合わせ
ピクセル正規化
差分処理
顔生体検出モデル
Ｎ－１フレームの第１の差分画像
時空間特徴抽出層 ４０１
第１の完全接続層 ４０２
第１の正規化層 ４０３
第１の検出結果
生体検出結果
第２の差分画像
画像特徴抽出層 ４０４
第２の完全接続層 ４０５
第２の正規化層 ４０６
第２の検出結果

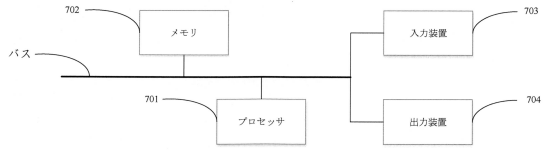
【図５】


顔生体検出装置
510
取得モジュール
520
第１の差分モジュール
530
第２の差分モジュール
540
第１の決定モジュール
550
第２の決定モジュール
551
判断ユニット
552
第１の決定ユニット
553
第２の決定ユニット

【図６】


顔生体検出装置
610
取得モジュール
620
第１の差分モジュール
630
第２の差分モジュール
640
第１の決定モジュール
650
第２の決定モジュール
651
判断ユニット
652
第１の決定ユニット
653
第２の決定ユニット
660
正規化モジュール

【図７】


バス
702
メモリ
701
プロセッサ
703
入力装置
704
出力装置

フロントページの続き

８５， Ｃｈｉｎａ
(74)代理人 110000578
名古屋国際弁理士法人
(72)発明者 ワン クヤオ
中華人民共和国 １０００８５ ベイジン ハイディエン ディストリクト シャンディー テンス ストリート ナンバー １０ バイドゥ キャンパス ２／エフ
審査官 淀川 滉也

(56)参考文献 特表２０１８－５０８８７５（ＪＰ，Ａ）
特開２００６－３３０９３６（ＪＰ，Ａ）
Emna FOURATI et al.，“ Face anti-spoofing with image quality assessment ”，2017 2nd International Conference on Bio-engineering for Smart Technologies (BioSMART)，2017年08月

(58)調査した分野 (Int.Cl.，ＤＢ名)
Ｇ０６Ｖ ４０／４０
Ｇ０６Ｔ ７／２０