

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6905601号  
(P6905601)

(45) 発行日 令和3年7月21日(2021.7.21)

(24) 登録日 令和3年6月29日(2021.6.29)

(51) Int.Cl.

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 6 6 0 B

請求項の数 12 (全 18 頁)

(21) 出願番号	特願2019-570048 (P2019-570048)	(73) 特許権者	516158932
(86) (22) 出願日	平成30年11月21日(2018.11.21)		ベイジン センスタ임 テクノロジー
(65) 公表番号	特表2020-525901 (P2020-525901A)		デベロップメント カンパニー, リミテ
(43) 公表日	令和2年8月27日(2020.8.27)		ッド
(86) 国際出願番号	PCT/CN2018/116600		中華人民共和国 ベイジン 100080
(87) 国際公開番号	W02019/153830		, ハイディアン, ナンバー 58 ノ
(87) 国際公開日	令和1年8月15日(2019.8.15)		ースウェスト 4ティーエイチ リング
審査請求日	令和1年12月18日(2019.12.18)		ロード, フロア 11ティーエイチ,
(31) 優先権主張番号	201810145717.3		ルーム 1101-1117
(32) 優先日	平成30年2月12日(2018.2.12)	(74) 代理人	110001427
(33) 優先権主張国・地域又は機関	中国 (CN)		特許業務法人前田特許事務所

最終頁に続く

(54) 【発明の名称】 歩行者再認識方法、装置、電子機器および記憶媒体

(57) 【特許請求の範囲】

【請求項 1】

目標歩行者を含む目標映像および少なくとも一つの候補映像を取得することと、

前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化することと、

符号化結果に基づいて、前記目標映像セグメントにおける歩行者と前記候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定することと、

前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うことと、を含み、

前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化することは、

各前記目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを取得し、各前記候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得することと、

前記インデックス特徴ベクトル、前記第一目標特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて注意重みベクトルを生成することと、

前記注意重みベクトル、前記第二目標特徴ベクトルおよび前記第二候補特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果および各前記候補映像セグメントの符号

10

20

化結果を得ることと、を含む、  
歩行者再認識方法。

【請求項 2】

各前記目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを取得し、各前記候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得することは、

各前記目標映像フレームの画像特徴ベクトルおよび各前記候補映像フレームの画像特徴ベクトルをそれぞれ抽出することと、

各前記目標映像フレームの画像特徴ベクトルに基づいて各前記目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを生成し、各前記候補映像フレームの画像特徴ベクトルに基づいて各前記候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを生成することと、を含む、請求項 1 に記載の方法。

10

【請求項 3】

前記インデックス特徴ベクトル、前記第一目標特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて注意重みベクトルを生成することは、

前記インデックス特徴ベクトルおよび前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標注意重みベクトルを生成し、前記インデックス特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補注意重みベクトルを生成することを含む、請求項 1 または 2 に記載の方法。

20

【請求項 4】

前記インデックス特徴ベクトルおよび前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標注意重みベクトルを生成することは、

前記インデックス特徴ベクトル、各前記目標映像フレームの前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標ヒートマップを生成することと、

前記目標ヒートマップを正規化して各前記目標映像フレームの目標注意重みベクトルを得ることと、を含む、

および / または、

前記インデックス特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補注意重みベクトルを生成することは、

30

前記インデックス特徴ベクトル、各前記候補映像フレームの前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補ヒートマップを生成することと、

前記候補ヒートマップを正規化して各前記候補映像フレームの候補注意重みベクトルを得ることと、を含む、請求項 3 に記載の方法。

【請求項 5】

前記注意重みベクトル、前記第二目標特徴ベクトルおよび前記第二候補特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果および各前記候補映像セグメントの符号化結果を得ることは、

各前記目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果を得て、各前記候補映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各前記候補映像セグメントの符号化結果を得ることを含む、請求項 1 から 4 のいずれか一項に記載の方法。

40

【請求項 6】

各前記目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果を得ることは、

各前記目標映像フレームの目標注意重みベクトルを各々の目標映像フレームの第二目標特徴ベクトルと乗算することと、

各前記目標映像フレームの乗算結果を時間次元で加算し、各前記目標映像セグメントの符号化結果を得ることと、を含む、

50

および/または、

各前記目標映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各前記候補映像セグメントの符号化結果を得ることは、

各前記候補映像フレームの候補注意重みベクトルを各々の候補映像フレームの第二候補特徴ベクトルと乗算することと、

各前記候補映像フレームの乗算結果を時間次元で加算し、各前記候補映像セグメントの符号化結果を得ることと、を含む、請求項 5 に記載の方法。

【請求項 7】

符号化結果に基づいて各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定することは、

各前記目標映像セグメントの符号化結果と各前記候補映像セグメントの符号化結果を順に減算することと、

減算操作の結果を各次元で二乗することと、

二乗操作によって得られた特徴ベクトルを完全接続して二次元の特徴ベクトルを得ることと、

前記二次元の特徴ベクトルを正規化し、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを得ることと、を含む、請求項 1 から 6 のいずれか一項に記載の方法。

【請求項 8】

前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うことは、

少なくとも一つの前記候補映像内の各前記候補映像セグメントについて、比率閾値が予め設定されたスコアが最も高い前記類似度スコアを、各前記候補映像の類似度スコアとして加算することと、

各前記候補映像の類似度スコアを降順に配列することと、

前に配列された一つまたはいくつかの前記候補映像を前記目標映像と同一目標歩行者を含む映像として決定することと、を含む、請求項 1 から 7 のいずれか一項に記載の方法。

【請求項 9】

目標歩行者を含む目標映像および少なくとも一つの前記候補映像を取得するように構成された取得モジュールと、

前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化するように構成された符号化モジュールと、

符号化結果に基づいて、前記目標映像セグメントにおける歩行者と前記候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定するように構成された決定モジュールと、

前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うように構成された認識モジュールと、を含む、

前記符号化モジュールは、

各前記目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを取得し、各前記候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得するように構成された特徴ベクトル取得モジュールと、

前記インデックス特徴ベクトル、前記第一目標特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて注意重みベクトルを生成するように構成された重みベクトル生成モジュールと、

前記注意重みベクトル、前記第二目標特徴ベクトルおよび前記第二候補特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果および各前記候補映像セグメントの符号化結果を得るように構成された符号化結果取得モジュールと、を含む、歩行者再認識装置

10

20

30

40

50

。

【請求項 10】

プロセッサ、および

前記プロセッサに請求項 1 から 8 のいずれか一項に記載の歩行者再認識方法を実行させる少なくとも一つの実行可能命令を記憶するためのメモリを含む、電子機器。

【請求項 11】

プロセッサに実行される時に請求項 1 から 8 のいずれか一項に記載の歩行者再認識方法を実現するコンピュータプログラムが記憶されている、コンピュータ読み取り可能記憶媒体。

【請求項 12】

プロセッサに実行される時に請求項 1 から 8 のいずれか一項に記載の歩行者再認識方法を実現するための少なくとも一つの実行可能命令を含む、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願の相互参照)

本願は 2018 年 2 月 12 日に出願された、出願番号 201810145717.3 の中国特許出願に基づいて提出され、該中国特許出願の優先権を主張し、その開示の全てが参照によって本願に組み込まれる。

【0002】

本願の実施例は画像処理の技術分野に関し、特に歩行者再認識方法、装置、電子機器および記憶媒体に関する。

【背景技術】

【0003】

歩行者再認識はスマート映像監視システムにおける重要な技術の一つであり、その目的は指定された目標映像と候補映像との間の類似度を測定し、さらに大量の候補映像から目標映像と同一歩行者を含む候補映像を探し出すことである。

【0004】

従来の歩行者再認識方法は主に一つの完全な映像を符号化し、符号化結果により全目標映像と全候補映像との間の類似度を測定し、歩行者再認識の効果が低い。

【発明の概要】

【0005】

本願の実施例は歩行者再認識の技術的解決手段を提供する。

【0006】

本願の実施例の第一態様によれば、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得することと、前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化することと、符号化結果に基づいて、前記目標映像セグメントにおける歩行者と前記候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定することと、前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うことと、を含む歩行者再認識方法が提供される。

【0007】

一実施例では、前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化することは、各前記目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを取得し、各前記候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得することと、前記インデックス特徴ベクトル、前記第一目標特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて注意重みベクトルを生成することと、前記注意重み

10

20

30

40

50

ベクトル、前記第二目標特徴ベクトルおよび前記第二候補特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果および各前記候補映像セグメントの符号化結果を得ることと、を含む。

【0008】

一実施例では、各前記目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを取得し、各前記候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得することは、各前記目標映像フレームの画像特徴ベクトルおよび各前記候補映像フレームの画像特徴ベクトルをそれぞれ抽出することと、各前記目標映像フレームの画像特徴ベクトルに基づいて各前記目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを生成し、各前記候補映像フレームの画像特徴ベクトルに基づいて各前記候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを生成することと、を含む。

10

【0009】

一実施例では、前記インデックス特徴ベクトル、前記第一目標特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて注意重みベクトルを生成することは、前記インデックス特徴ベクトルおよび前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標注意重みベクトルを生成し、前記インデックス特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補注意重みベクトルを生成することを含む。

20

【0010】

一実施例では、前記インデックス特徴ベクトルおよび前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標注意重みベクトルを生成することは、前記インデックス特徴ベクトル、各前記目標映像フレームの前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標ヒートマップを生成することと、前記目標ヒートマップを正規化して各前記目標映像フレームの目標注意重みベクトルを得ることと、を含み、および/または、前記インデックス特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補注意重みベクトルを生成することは、前記インデックス特徴ベクトル、各前記候補映像フレームの前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補ヒートマップを生成することと、前記候補ヒートマップを正規化して各前記候補映像フレームの候補注意重みベクトルを得ることと、を含む。

30

【0011】

一実施例では、前記注意重みベクトル、前記第二目標特徴ベクトルおよび前記第二候補特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果および各前記候補映像セグメントの符号化結果を得ることは、各前記目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果を得て、各前記候補映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各前記候補映像セグメントの符号化結果を得ることを含む。

【0012】

一実施例では、各前記目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果を得ることは、各前記目標映像フレームの目標注意重みベクトルを各々の目標映像フレームの第二目標特徴ベクトルと乗算することと、各前記目標映像フレームの乗算結果を時間次元で加算し、各前記目標映像セグメントの符号化結果を得ることと、を含み、および/または、各前記目標映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各前記候補映像セグメントの符号化結果を得ることは、各前記候補映像フレームの候補注意重みベクトルを各々の候補映像フレームの第二候補特徴ベクトルと乗算することと、各前記候補映像フレームの乗算結果を時間次元で加算し、各前記候補映像セグメントの符号化結果を得ることと、を含む。

40

【0013】

50

－実施例では、符号化結果に基づいて各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定することは、各前記目標映像セグメントの符号化結果と各前記候補映像セグメントの符号化結果を順に減算することと、減算操作の結果を各次元で二乗することと、二乗操作によって得られた特徴ベクトルを完全接続して二次元の特徴ベクトルを得ることと、前記二次元の特徴ベクトルを正規化し、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを得ることと、を含む。

【0014】

－実施例では、前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うことは、少なくとも一つの前記候補映像内の各前記候補映像セグメントについて、比率閾値が予め設定されたスコアが最も高い前記類似度スコアを、各前記候補映像の類似度スコアとして加算することと、各前記候補映像の類似度スコアを降順に配列することと、前に配列された一つまたはいくつかの前記候補映像を前記目標映像と同一目標歩行者を含む映像として決定することと、を含む。

10

【0015】

本願の実施例の第二態様によれば、目標歩行者を含む目標映像および少なくとも一つの前記候補映像を取得するように構成された取得モジュールと、前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化するように構成された符号化モジュールと、符号化結果に基づいて、前記目標映像セグメントにおける歩行者と前記候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定するように構成された決定モジュールと、前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うように構成された認識モジュールと、を含む歩行者再認識装置が提供される。

20

【0016】

－実施例では、前記符号化モジュールは、各前記目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを取得し、各前記候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得するように構成された特徴ベクトル取得モジュールと、前記インデックス特徴ベクトル、前記第一目標特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて注意重みベクトルを生成するように構成された重みベクトル生成モジュールと、前記注意重みベクトル、前記第二目標特徴ベクトルおよび前記第二候補特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果および各前記候補映像セグメントの符号化結果を得るように構成された符号化結果取得モジュールと、を含む。

30

【0017】

－実施例では、前記特徴ベクトル取得モジュールは、各前記目標映像フレームの画像特徴ベクトルおよび各前記候補映像フレームの画像特徴ベクトルをそれぞれ抽出し、そして各前記目標映像フレームの画像特徴ベクトルに基づいて各前記目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各前記目標映像セグメントのインデックス特徴ベクトルを生成し、各前記候補映像フレームの画像特徴ベクトルに基づいて各前記候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを生成するように構成される。

40

【0018】

－実施例では、前記重みベクトル生成モジュールは、前記インデックス特徴ベクトルおよび前記第一目標特徴ベクトルに基づいて各前記目標映像フレームの目標注意重みベクトルを生成し、前記インデックス特徴ベクトルおよび前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補注意重みベクトルを生成するように構成される。

【0019】

－実施例では、前記重みベクトル生成モジュールは、前記インデックス特徴ベクトル、各前記目標映像フレームの前記第一目標特徴ベクトルに基づいて各前記目標映像フレーム

50

の目標ヒートマップを生成し、そして前記目標ヒートマップを正規化して各前記目標映像フレームの目標注意重みベクトルを得るように、および/または、前記インデックス特徴ベクトル、各前記候補映像フレームの前記第一候補特徴ベクトルに基づいて各前記候補映像フレームの候補ヒートマップを生成し、そして前記候補ヒートマップを正規化して各前記候補映像フレームの候補注意重みベクトルを得るように構成される。

【0020】

一実施例では、前記符号化結果取得モジュールは、各前記目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各前記目標映像セグメントの符号化結果を得て、各前記候補映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各前記候補映像セグメントの符号化結果を得るように構成される。

10

【0021】

一実施例では、前記符号化結果取得モジュールは、各前記目標映像フレームの目標注意重みベクトルを各々の目標映像フレームの第二目標特徴ベクトルと乗算し、そして各前記目標映像フレームの乗算結果を時間次元で加算し、各前記目標映像セグメントの符号化結果を得るように、および/または、各前記候補映像フレームの候補注意重みベクトルを各々の候補映像フレームの第二候補特徴ベクトルと乗算し、そして各前記候補映像フレームの乗算結果を時間次元で加算し、各前記候補映像セグメントの符号化結果を得るように構成される。

【0022】

一実施例では、前記決定モジュールは、各前記目標映像セグメントの符号化結果と各前記候補映像セグメントの符号化結果を順に減算し、減算操作の結果を各次元で二乗し、二乗操作によって得られた特徴ベクトルを完全接続して二次元の特徴ベクトルを得て、そして前記二次元の特徴ベクトルを正規化し、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを得るように構成される。

20

【0023】

一実施例では、前記認識モジュールは、少なくとも一つの前記候補映像内の各前記候補映像セグメントについて、比率閾値が予め設定されたスコアが最も高い前記類似度スコアを、各前記候補映像の類似度スコアとして加算し、各前記候補映像の類似度スコアを降順に配列し、そして前に配列された一つまたはいくつかの前記候補映像を前記目標映像と同一目標歩行者を含む映像として決定するように構成される。

30

【0024】

本願の実施例の第三態様によれば、プロセッサ、および前記プロセッサに第一態様に記載の歩行者再認識方法を実行させる少なくとも一つの実行可能命令を記憶するためのメモリを含む電子機器が提供される。

【0025】

本願の実施例の第四態様によれば、プロセッサに実行される時に第一態様に記載の歩行者再認識方法を実現するコンピュータプログラムが記憶されているコンピュータ読み取り可能記憶媒体が提供される。

【0026】

本願の実施例の第五態様によれば、プロセッサに実行される時に第一態様に記載の歩行者再認識方法を実現するための少なくとも一つの実行可能命令を含むコンピュータプログラム製品が提供される。

40

【0027】

本願の実施例は歩行者再認識を行う時、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得し、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化し、符号化結果に基づいて各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定し、そして類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行う。映像セグメントに含まれるフレーム数は全映像に含まれるフレーム数より大幅に少ないため、映像セグメントにおける歩行者表面情報の変化度は全映像における歩行者表面情報の変化度より大幅に低い

50

。全目標映像および全候補映像に対する符号化に比べ、各目標映像セグメントおよび各候補映像セグメントに対する符号化は、歩行者表面情報の変化を効果的に低減し、同時に異なる映像フレーム内の歩行者表面情報の多様性および映像フレーム間の動的関連性により、歩行者表面情報の利用率を向上させ、各目標映像セグメントと各候補映像セグメントとの間の類似度スコア計算の符号化結果による正解率を向上させ、さらに歩行者再認識の正解率を向上可能である。

【図面の簡単な説明】

【0028】

【図1】本願の実施例に係る歩行者再認識方法の一実施例のフローチャートである。

【図2】本願の実施例に係る歩行者再認識方法の一実施例の計算フレームワーク模式図である。

10

【図3】本願の実施例に係る歩行者再認識方法の別の実施例のフローチャートである。

【図4】本願の実施例に係る歩行者再認識方法における注意符号化メカニズムの模式図である。

【図5】本願の実施例に係る歩行者再認識装置の一実施例の構成模式図である。

【図6】本願の実施例に係る歩行者再認識装置の別の実施例の構成模式図である。

【図7】本願の実施例に係る電子機器の一実施例の構成模式図である。

【発明を実施するための形態】

【0029】

以下に図面（複数の図面における同じ符号は同じ要素を表す）と実施例を関連付けて、本発明の実施例の具体的な実施形態をさらに詳細に説明する。以下の実施例は本発明を説明するためのものであるが、本発明の範囲を限定するものではない。

20

【0030】

当業者であれば、本発明の実施例における「第一」、「第二」などの用語は異なるステップ、機器またはモジュールなどを区別するためのものに過ぎず、なんらの特定の技術的意味も有さず、またそれらの間の必然的な論理的順序を表すものでもない。

【0031】

図1を参照すると、本願の実施例に係る歩行者再認識方法の一実施例のフローチャートが示される。

【0032】

30

本願の実施例の歩行者再認識方法はメモリに記憶された関連命令を電子機器のプロセッサによって呼び出して以下のステップを実行する。

【0033】

ステップS100において、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得する。

【0034】

本願の実施例における目標映像は一つ以上の目標歩行者を含んでもよく、候補映像は一つ以上の候補歩行者を含んでもよくまたは候補歩行者を含まなくてもよい。本願の実施例における目標映像および少なくとも一つの候補映像は映像収集機器に由来する映像画像であってもよく、また他の機器から由来してもよく、本願の実施例の一つの目的は少なくとも一つの候補映像から候補歩行者が目標歩行者と同一歩行者である候補映像を得ることである。

40

【0035】

任意選択的な一例では、該ステップS100はメモリに記憶された対応する命令をプロセッサによって呼び出して実行してもよいし、プロセッサにより運用される取得モジュール50によって実行してもよい。

【0036】

ステップS102において、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化する。

【0037】

50



まず、目標映像および候補映像のクリップを分割し、目標映像内の各目標映像セグメントおよび候補映像内の各候補映像セグメントを生成する。ここで、各目標映像セグメントは固定の長さ、各候補映像セグメントは固定の長さを有し、かつ、各目標映像セグメントの長さと各候補映像セグメントの長さは同じであってもでなくともよい。

【0038】

続いて、各目標映像セグメントおよび各候補映像セグメントをそれぞれ符号化し、各目標映像セグメントの符号化結果および各候補映像セグメントの符号化結果を得る。

【0039】

任意選択的な一例では、該ステップS102はメモリに記憶された対応する命令をプロセッサによって呼び出して実行してもよいし、プロセッサにより運用される符号化モジュール52によって実行してもよい。

10

【0040】

ステップS104において、符号化結果に基づいて各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定する。

【0041】

本願の実施例では、各目標映像セグメントの符号化結果は各目標映像セグメントにおける歩行者特徴ベクトルの一具現化形態、各候補映像セグメントの符号化結果は各候補映像セグメントにおける歩行者特徴ベクトルの一具現化形態と見なしてもよい。または、符号化結果は歩行者特徴ベクトルそのものである。ある目標映像セグメントとある候補映像セグメントとの間の歩行者特徴ベクトルが同じである場合または類似する場合、該目標映像セグメントと該候補映像セグメントとは同一目標歩行者を含む可能性が高いこと、即ち該目標映像セグメントと該候補映像セグメントとの間の類似度スコアが高いことが示され、ある目標映像セグメントとある候補映像セグメントとの間の歩行者特徴ベクトルが異なる場合、該目標映像セグメントと該候補映像セグメントとは同一目標歩行者を含む可能性が低いこと、即ち該目標映像セグメントと該候補映像セグメントとの間の類似度スコアが低いことが示される。

20

【0042】

任意選択的な一例では、該ステップS104はメモリに記憶された対応する命令をプロセッサによって呼び出して実行してもよいし、プロセッサにより運用される決定モジュール54によって実行してもよい。

30

【0043】

ステップS106において、類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行う。

【0044】

各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを得てから、類似度スコアに基づいて少なくとも一つの候補映像の類似度スコアを得ることができる。類似度スコアが高い候補映像を目標映像と同一目標歩行者を含む候補映像として決定する。

【0045】

任意選択的な一例では、該ステップS106はメモリに記憶された対応する命令をプロセッサによって呼び出して実行してもよいし、プロセッサにより運用される認識モジュール56によって実行してもよい。

40

【0046】

本願の実施例が提供する歩行者再認識方法は図2に示す計算フレームワークで実行可能である。まず、映像（目標映像および少なくとも一つの候補映像を含む）を分割し、固定の長さを有する映像セグメントを生成する。図中、 $p$ は目標映像を表し、 $g$ は少なくとも一つの候補映像のうちの一つの候補映像を表し、 $p_n$ は目標映像 $p$ 内の一つの目標映像セグメントであり、 $g_k$ は候補映像 $g$ 内の一つの候補映像セグメントである。目標映像 $p$ と候補映像 $g$ 内の任意の二つの映像セグメント間の類似度を測定するために、相乗注意メカニズムを有する深層ネットワークを使用する。該深層ネットワークは目標映像セグメント $p_n$ および候補映像セグメント $g_k$ を入力項目とし、出力項目 $m(p_n, g_k)$ は目標映

50

像セグメント  $p_n$  と候補映像セグメント  $g_k$  との間の類似度スコアとする。目標映像  $p$  および候補映像  $g$  内の二つの映像セグメント（目標映像セグメントおよび候補映像セグメント）につき、複数の映像セグメント間の類似度スコアを得ることができる。目標映像  $p$  と候補映像  $g$  との間の類似度を有効に推定するために、競争的メカニズムを用いて類似度が高い一部の類似度スコアを選択し、これらの類似度スコアを加算することで目標映像  $p$  と候補映像  $g$  との間の類似度の確実な推定  $c(p, g)$  を得るようにしてもよい。

【0047】

本願の実施例は歩行者再認識を行う時、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得し、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化し、符号化結果に基づいて各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定し、そして類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行う。映像セグメントに含まれるフレーム数は全映像に含まれるフレーム数より大幅に少ないため、映像セグメントにおける歩行者表面情報の変化度は全映像における歩行者表面情報の変化度より大幅に低い。全目標映像および全候補映像に対する符号化に比べ、各目標映像セグメントおよび各候補映像セグメントに対する符号化は、歩行者表面情報の変化を効果的に低減し、同時に異なる映像フレーム内の歩行者表面情報の多様性および映像フレーム間の動的関連性により、歩行者表面情報の利用率を向上させ、各目標映像セグメントと各候補映像セグメントとの間の類似度スコア計算の符号化結果による正解率を向上させ、さらに歩行者再認識の正解率を向上可能である。

【0048】

図3を参照すると、本願の実施例に係る歩行者再認識方法の別の実施例のフローチャートが示される。

【0049】

説明すべきは、本願の各実施例に記述した部分は着目している点が異なり、ある実施例に詳細に記述されていない部分は本出願の他の実施例での記述および説明を参照すればよく、詳細な説明は繰り返さないということである。

【0050】

ステップS300において、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得する。

【0051】

ステップS302において、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化する。

【0052】

任意選択的に、本ステップS302は以下のステップを含んでもよい。

【0053】

ステップS3020において、各目標映像セグメント内の各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各目標映像セグメントのインデックス特徴ベクトルを取得し、各候補映像セグメント内の各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得する。

【0054】

任意選択的な一実施形態では、ニューラルネットワークによって、映像フレームにおける画像特徴、例えば歩行者特徴、背景特徴などを反映するための、各目標映像フレームの画像特徴ベクトルおよび各候補映像フレームの画像特徴ベクトルを抽出してもよい。目標映像フレームについて、各目標映像フレームの画像特徴ベクトルに基づいて各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトル、ならびに目標映像セグメントの情報を含んで有用情報とノイズ情報を効果的に区別可能な各目標映像セグメントのインデックス特徴ベクトルを生成する。候補映像フレームについて、各候補映像フレームの画像特徴ベクトルに基づいて各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを生成する。具体的には、フレーム毎の特徴の線形変換によって第一目標特

徴ベクトル（「キー」特徴ベクトル）および第一候補特徴ベクトル（「キー」特徴ベクトル）を生成し、フレーム毎の特徴の別の線形変換によって第二目標特徴ベクトル（「値」特徴ベクトル）および第二候補特徴ベクトル（「値」特徴ベクトル）を生成し、長期短期記憶（Long Short-Term Memory、LSTM）ネットワークおよび各目標映像セグメントの各目標映像フレームの画像特徴ベクトルによって、目標映像セグメントから生成され目標映像セグメント自身および全ての候補映像セグメントに作用する各目標映像セグメントのインデックス特徴ベクトルを生成するようにしてもよい。

【0055】

ステップS3022において、インデックス特徴ベクトル、第一目標特徴ベクトルおよび第一候補特徴ベクトルに基づいて注意重みベクトルを生成する。

10

【0056】

本願の実施例では、第一目標特徴ベクトルおよび第一候補特徴ベクトルは注意重みベクトルを生成するために用いられる。任意選択的な一実施形態では、目標映像フレームについて、インデックス特徴ベクトルおよび第一目標特徴ベクトルに基づいて各目標映像フレームの目標注意重みベクトルを生成してもよく、任意選択的に、インデックス特徴ベクトル、各目標映像フレームの第一目標特徴ベクトルに基づいて各目標映像フレームの目標ヒートマップを生成し、具体的には、インデックス特徴ベクトル、各目標映像フレームの第一目標特徴ベクトルに基づいて内積を行って各目標映像フレームの目標ヒートマップを得て、そして時間次元でsoftmax関数によって目標ヒートマップを正規化して各目標映像フレームの目標注意重みベクトルを得る。候補映像フレームについて、インデックス特徴ベクトルおよび第一候補特徴ベクトルに基づいて各候補映像フレームの候補注意重みベクトルを生成してもよく、任意選択的に、インデックス特徴ベクトル、各候補映像フレームの第一候補特徴ベクトルに基づいて各候補映像フレームの候補ヒートマップを生成し、具体的には、インデックス特徴ベクトル、各候補映像フレームの第一候補特徴ベクトルに基づいて内積を行って各候補映像フレームの候補ヒートマップを得て、そして時間次元でsoftmax関数によって候補ヒートマップを正規化して各候補映像フレームの候補注意重みベクトルを得る。

20

【0057】

注意重みベクトルは符号化プロセスで有効な歩行者特徴を増強するための、識別力情報を有する重みベクトルであり、ノイズ情報による影響を軽減可能である。

30

【0058】

ステップS3024において、注意重みベクトル、第二目標特徴ベクトルおよび第二候補特徴ベクトルに基づいて各目標映像セグメントの符号化結果および候補映像セグメントの符号化結果を得る。

【0059】

本願の実施例では、第二目標特徴ベクトルは目標映像セグメント内の各フレームの画像特徴を反映するために用いられ、第二候補特徴ベクトルは候補映像セグメント内の各フレームの画像特徴を反映するために用いられる。任意選択的な一実施形態では、目標映像フレームについて、各目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各目標映像セグメントの符号化結果を得る。具体的には、各目標映像フレームの目標注意重みベクトルを各々の目標映像フレームの第二目標特徴ベクトルと乗算し、そして各目標映像フレームの乗算結果を時間次元で加算し、各目標映像セグメントの符号化結果を得る。候補映像フレームについて、各候補映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各候補映像セグメントの符号化結果を得る。任意選択的に、各候補映像フレームの候補注意重みベクトルを各々の候補映像フレームの第二候補特徴ベクトルと乗算し、そして各候補映像フレームの乗算結果を時間次元で加算し、各候補映像セグメントの符号化結果を得る。

40

【0060】

本願の実施例のステップS302は注意符号化メカニズムによって実現可能、即ち映像セグメント（目標映像セグメントおよび候補映像セグメント）における異なるフレーム特

50

徴の抽出によって映像セグメントの符号化結果を得ることが可能であり、そのプロセスは図4に示されたとおりである。まず目標映像セグメントにおける各目標映像フレームおよび候補映像セグメントにおける各候補映像フレームについて畳み込みニューラルネットワーク特徴を抽出し、畳み込みニューラルネットワーク特徴に基づいて各目標映像フレームまたは各候補映像フレームにそれぞれ対応する「キー」特徴ベクトルおよび「値」特徴ベクトルを生成し、各目標映像フレームまたは各候補映像フレームの「キー」特徴ベクトルを各目標映像セグメントのインデックス特徴ベクトルと内積してヒートマップを形成させ、ヒートマップによって目標映像フレームまたは候補映像フレーム内の各特徴とグローバル情報との関連性を反映させる。ヒートマップを時間次元でsoftmax関数によって正規化して注意重みベクトルを形成させ、該注意重みベクトルを各映像フレームの「値」特徴ベクトルに対応して各次元でそれぞれと乗算し、異なる映像フレームによって得られた結果を時間次元で加算し、さらに各映像セグメントの符号化結果を得る。

10

**【0061】**

ステップS304において、符号化結果に基づいて各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定する。

**【0062】**

任意選択的な一実施形態では、各目標映像セグメントの符号化結果と各候補映像セグメントの符号化結果を順に減算、二乗、完全接続および正規化し、各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを得る。具体的には、各目標映像セグメントの符号化結果と各候補映像セグメントの符号化結果を順に減算し、続いて各画像次元で二乗する。ここで、画像次元は、頭部画像次元、上半身画像次元、下半身画像次元などを含む歩行者画像次元、および、建築物画像次元、街路画像次元などを含む背景画像次元を含むが、これらに限定されない。二乗操作によって得られた特徴ベクトルを完全接続層によって二次元の特徴ベクトルを得て、最後にsigmoid関数による非線形正規化によって各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを得る。

20

**【0063】**

ステップS306において、類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行う。

**【0064】**

任意選択的な一実施形態では、少なくとも一つの候補映像のうちの各候補映像について、予め設定された閾値以上の類似度スコアまたはスコアが高い類似度スコア（例えば、前20%に配列された類似度スコア）を、各候補映像の類似度スコアとして加算し、各候補映像の類似度スコアを降順に配列し、そして前に配列された一つまたはいくつかの候補映像を目標映像と同一目標歩行者を含む映像として決定する。ここで、予め設定された閾値は実際の状況に応じて設定してもよく、スコアが高いとは相対的に高いということである。

30

**【0065】**

本願の実施例は歩行者再認識を行う時、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得し、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化し、符号化結果に基づいて各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定し、そして類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行う。映像セグメントに含まれるフレーム数は全映像に含まれるフレーム数より大幅に少ないため、映像セグメントにおける歩行者表面情報の変化度は全映像における歩行者表面情報の変化度より大幅に低い。全目標映像および全候補映像に対する符号化に比べ、各目標映像セグメントおよび各候補映像セグメントに対する符号化は、歩行者表面情報の変化を効果的に低減し、同時に異なる映像フレーム内の歩行者表面情報の多様性および映像フレーム間の動的関連性により、歩行者表面情報の利用率を向上させ、各目標映像セグメントと各候補映像セグメントとの間の類似度スコア計算の符号化結果による正解率を向上させ、さらに歩行者再認識の正解率を向上可能である。

40

50

## 【 0 0 6 6 】

本願の実施例における候補映像の符号化結果は目標映像セグメントのインデックス特徴ベクトルおよび候補映像セグメントの「キー」特徴ベクトルによって得られ、符号化プロセスで、目標映像セグメントのインデックス特徴ベクトルを指導情報として用い、類似度スコア決定の候補映像符号化結果による正確性を向上させる。目標映像セグメントのインデックス特徴ベクトルにより各候補映像フレームの注意重みベクトルを推定し、候補映像における異常候補映像フレームが候補映像セグメントの符号化結果に与える影響を軽減し、候補映像における歩行者再認識の指向性を向上させる。

## 【 0 0 6 7 】

本願の実施例は目標映像および候補映像をクリップに分割し、目標映像セグメントおよび候補映像セグメントを符号化し、候補映像における歩行者が一部の候補映像フレームにおいて遮蔽された場合、類似度スコアが高い候補映像セグメントを候補映像の有効候補映像セグメントとして選択し、類似度スコアが低い候補映像セグメントを無視する。

10

## 【 0 0 6 8 】

図 5 を参照すると、本願の実施例に係る歩行者再認識装置の一実施例の構成模式図が示される。

## 【 0 0 6 9 】

本願の実施例が提供する歩行者再認識装置は、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得するように構成された取得モジュール 5 0 と、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化するように構成された符号化モジュール 5 2 と、符号化結果に基づいて、目標映像セグメントにおける歩行者と候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定するように構成された決定モジュール 5 4 と、類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行うように構成された認識モジュール 5 6 と、を含む。

20

## 【 0 0 7 0 】

本願の実施例の歩行者再認識装置は上記実施例における対応する歩行者再認識方法を実現するためのものであり、かつ対応する方法の実施例の有益な効果を有し、ここでは説明を省略する。

30

## 【 0 0 7 1 】

図 6 を参照すると、本願の実施例に係る歩行者再認識装置の別の実施例の構成模式図が示される。

## 【 0 0 7 2 】

本願の実施例が提供する歩行者再認識装置は、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得するように構成された取得モジュール 6 0 と、目標映像内の各目標映像セグメントおよび少なくとも一つの候補映像内の各候補映像セグメントをそれぞれ符号化するように構成された符号化モジュール 6 2 と、符号化結果に基づいて、目標映像セグメントにおける歩行者と候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを決定するように構成された決定モジュール 6 4 と、類似度スコアに基づいて少なくとも一つの候補映像における歩行者再認識を行うように構成された認識モジュール 6 6 と、を含む。

40

## 【 0 0 7 3 】

任意選択的に、符号化モジュール 6 2 は、各目標映像セグメントにおける各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各目標映像セグメントのインデックス特徴ベクトルを取得し、各候補映像セグメントにおける各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを取得するように構成された特徴ベクトル取得モジュール 6 2 0 と、インデックス特徴ベクトル、第一目標特徴ベクトルおよび第一候補特徴ベクトルに基づいて注意重みベクトルを生成するように構成された重

50

みベクトル生成モジュール622と、注意重みベクトル、第二目標特徴ベクトルおよび第二候補特徴ベクトルに基づいて各目標映像セグメントの符号化結果および各候補映像セグメントの符号化結果を得るように構成された符号化結果取得モジュール624と、を含む。

【0074】

任意選択的に、特徴ベクトル取得モジュール620は、各目標映像フレームの画像特徴ベクトルおよび各候補映像フレームの画像特徴ベクトルをそれぞれ抽出し、そして各目標映像フレームの画像特徴ベクトルに基づいて各目標映像フレームの第一目標特徴ベクトルおよび第二目標特徴ベクトルならびに各目標映像セグメントのインデックス特徴ベクトルを生成し、各候補映像フレームの画像特徴ベクトルに基づいて各候補映像フレームの第一候補特徴ベクトルおよび第二候補特徴ベクトルを生成するように構成される。

10

【0075】

任意選択的に、重みベクトル生成モジュール622は、インデックス特徴ベクトルおよび第一目標特徴ベクトルに基づいて各目標映像フレームの目標注意重みベクトルを生成し、インデックス特徴ベクトルおよび第一候補特徴ベクトルに基づいて各候補映像フレームの候補注意重みベクトルを生成するように構成される。

【0076】

任意選択的に、重みベクトル生成モジュール622は、インデックス特徴ベクトル、各目標映像フレームの第一目標特徴ベクトルに基づいて各目標映像フレームの目標ヒートマップを生成し、そして目標ヒートマップを正規化して各目標映像フレームの目標注意重みベクトルを得るように、および/または、インデックス特徴ベクトル、各候補映像フレームの第一候補特徴ベクトルに基づいて各候補映像フレームの候補ヒートマップを生成し、そして候補ヒートマップを正規化して各候補映像フレームの候補注意重みベクトルを得るように構成される。

20

【0077】

任意選択的に、符号化結果取得モジュール624は、各目標映像フレームの目標注意重みベクトルおよび第二目標特徴ベクトルに基づいて各目標映像セグメントの符号化結果を得て、各候補映像フレームの候補注意重みベクトルおよび第二候補特徴ベクトルに基づいて各候補映像セグメントの符号化結果を得るように構成される。

【0078】

30

任意選択的に、符号化結果取得モジュール624は、各目標映像フレームの目標注意重みベクトルを各々の目標映像フレームの第二目標特徴ベクトルと乗算し、そして各目標映像フレームの乗算結果を時間次元で加算し、各目標映像セグメントの符号化結果を得るように、および/または、各候補映像フレームの候補注意重みベクトルを各々の候補映像フレームの第二候補特徴ベクトルと乗算し、そして各候補映像フレームの乗算結果を時間次元で加算し、各候補映像セグメントの符号化結果を得るように構成される。

【0079】

任意選択的に、決定モジュール64は、各目標映像セグメントの符号化結果と各候補映像セグメントの符号化結果を順に減算し、減算操作の結果を各次元で二乗し、二乗操作によって得られた特徴ベクトルを完全接続して二次元の特徴ベクトルを得て、そして二次元の特徴ベクトルを正規化し、各目標映像セグメントと各候補映像セグメントとの間の類似度スコアを得るように構成される。

40

【0080】

任意選択的に、認識モジュール66は、少なくとも一つの候補映像内の各候補映像セグメントについて、比率閾値が予め設定されたスコアが最も高い類似度スコアを、各候補映像の類似度スコアとして加算し、各候補映像の類似度スコアを降順に配列し、そして前に配列された一つまたはいくつかの候補映像を目標映像と同一目標歩行者を含む映像として決定するように構成される。

【0081】

本願の実施例の歩行者再認識装置は上記実施例における対応する歩行者再認識方法を実

50

現するためのものであり、かつ対応する方法の実施例の有益な効果を有し、ここでは説明を省略する。

#### 【0082】

本願の実施例はさらに電子機器を提供し、例えば移動端末、パーソナルコンピュータ（PC）、タブレット、サーバなどであってもよい。以下に図7を参照すると、本願の実施例に係る歩行者再認識装置の実現に適する電子機器700の構成模式図が示される。図7に示すように、電子機器700はメモリおよびプロセッサを含んでもよい。具体的には、電子機器700は一つ以上のプロセッサ、通信要素などを含み、前記一つ以上のプロセッサは例えば、一つ以上の中央処理装置（CPU）701、および/または一つ以上の画像処理装置（GPU）713などであり、プロセッサは読み取り専用メモリ（ROM）702に記憶されている実行可能命令または記憶部分708からランダムアクセスメモリ（RAM）703にロードされた実行可能命令に従って様々な適当の動作および処理を実行できる。通信要素は通信コンポーネント712および/または通信インタフェース709を含む。ここで、通信コンポーネント712はネットワークカードを含むことができるが、これに限定されず、前記ネットワークカードはIB（Infiniband）ネットワークカードを含むことができるが、これに限定されず、通信インタフェース709はローカルエリアネットワーク（LAN、Local Area Network）カード、モデムのネットワークインタフェースカードなどの通信インタフェースを含み、かつインターネットなどのネットワークによって通信処理を実行する。

#### 【0083】

プロセッサは読み取り専用メモリ702および/またはランダムアクセスメモリ703と通信して実行可能命令を実行し、通信バス704を介して通信コンポーネント712と接続し、通信コンポーネント712によって他の目標機器と通信し、それにより本願の実施例が提供するいずれか一項の方法に対応する動作、例えば、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得することと、前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化することと、符号化結果に基づいて、前記目標映像セグメントにおける歩行者と前記候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定することと、前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うことと、を完了することができる。

#### 【0084】

また、RAM703には、装置の動作に必要な種々のプログラムおよびデータを記憶することができる。CPU701またはGPU713、ROM702およびRAM703は通信バス704を介して互いに接続される。RAM703が存在する場合、ROM702は任意選択的なモジュールとなる。RAM703は実行可能命令を記憶するか、または動作時にROM702へ実行可能命令を書き込み、実行可能命令によってプロセッサは上記通信方法に対応する動作を実行する。入力/出力（I/O）インタフェース705も通信バス704に接続される。通信コンポーネント712は統合設置してもよいし、また複数のサブモジュール（例えば複数のIBネットワークカード）を有するように設置してもよく、かつバスリンクに存在する。

#### 【0085】

キーボード、マウスなどを含む入力部分706、陰極線管（CRT）、液晶ディスプレイ（LCD）などおよびスピーカーなどを含む出力部分707、ハードディスクなどを含む記憶部分708、およびLANカード、モデムのネットワークインタフェースカードなどを含む通信インタフェース709といった部品は、I/Oインタフェース705に接続される。ドライバ710も必要に応じてI/Oインタフェース705に接続される。取り外し可能な媒体711、例えば磁気ディスク、光ディスク、磁気光ディスク、半導体メモリなどは、必要に応じてドライバ710に取り付けられ、それによってそこから読み出されたコンピュータプログラムが必要に応じて記憶部分708にインストールされる。

## 【 0 0 8 6 】

説明すべきは、図 7 に示すアーキテクチャは任意選択的な一実施形態に過ぎず、具体的な実践では、実際の必要に応じて上記図 7 の部品数およびタイプを選択、減少、増加または交換することができ、異なる機能部品の設置上でも、分離設置または統合設置などの実施形態を採用でき、例えば GPU と CPU は分離設置するかまたは GPU を CPU に統合するようにしてもよく、通信要素は分離設置してもよいし、また CPU または GPU に統合してもよいなどである。これらの置換可能な実施形態はいずれも本発明の保護範囲に属する。

## 【 0 0 8 7 】

本願の実施例に係る電子機器は上記実施例における対応する歩行者再認識方法を実現するために用いることができ、該電子機器内の各デバイスは上記方法の実施例における各ステップを実行するために用いることができ、例えば、上述した歩行者再認識方法はメモリに記憶された関連命令を電子機器のプロセッサによって呼び出すことで実現可能であり、明細書を簡潔にするために、ここでは説明を省略する。

10

## 【 0 0 8 8 】

本願の実施例によれば、フローチャートを参照しながら上述したプロセスはコンピュータプログラム製品として実現可能である。例えば、本願の実施例はコンピュータプログラム製品を含み、それは機械可読媒体に有形に具現化された、フローチャートに示す方法を実行するためのプログラムコードを含むコンピュータプログラムを含み、プログラムコードは本願の実施例により提供される方法を実行するステップに対応する命令、例えば、目標歩行者を含む目標映像および少なくとも一つの候補映像を取得することと、前記目標映像内の各目標映像セグメントおよび少なくとも一つの前記候補映像内の各候補映像セグメントをそれぞれ符号化することと、符号化結果に基づいて、前記目標映像セグメントにおける歩行者と前記候補映像セグメントにおける歩行者との特徴の類似度を特徴付けるための、各前記目標映像セグメントと各前記候補映像セグメントとの間の類似度スコアを決定することと、前記類似度スコアに基づいて少なくとも一つの前記候補映像における歩行者再認識を行うことと、を含むことができる。このような実施例では、該コンピュータプログラムは通信要素によってネットワークからダウンロードおよびインストールでき、および/または取り外し可能な媒体 7 1 1 からインストールできる。該コンピュータプログラムはプロセッサに実行される時、本願の実施例の方法に開示された機能を実行する。

20

30

## 【 0 0 8 9 】

本願の実施例の方法および装置、電子機器および記憶媒体は様々な方式で実現し得る。例えば、ソフトウェア、ハードウェア、ファームウェアまたはソフトウェア、ハードウェア、ファームウェアの任意の組み合わせで本願の実施例の方法および装置、電子機器および記憶媒体を実現できる。前記方法のステップに付けられる上記順序は説明するためのものに過ぎず、本願の実施例の方法のステップは、特に断らない限り、以上に具体的に記述した順序に限定されない。また、いくつかの実施例では、本願を記録媒体に記録されたプログラムとして実施してもよく、これらのプログラムは本願の実施例に係る方法を実現するための機械可読命令を含む。従って、本願は本願の実施例に係る方法を実行するためのプログラムを記憶する記録媒体をも包含する。

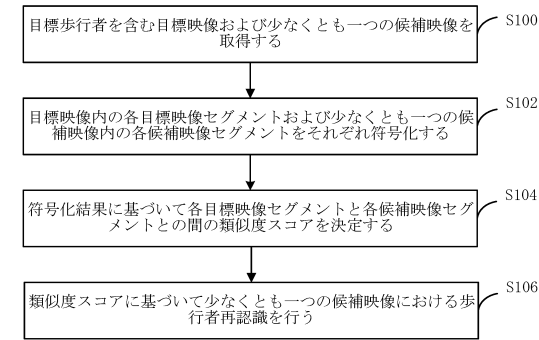
40

## 【 0 0 9 0 】

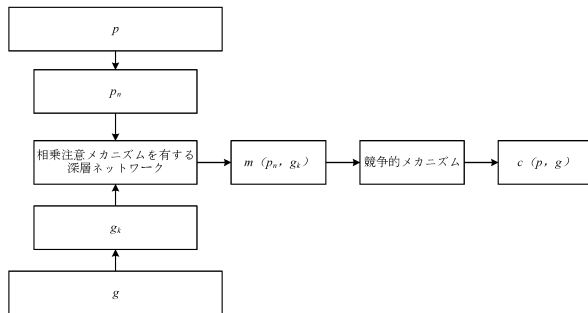
本願の実施例に対する記述は例示および説明のためのもので、漏れがないものまたは開示した形式に本発明を限定するものではなく、様々な修正および変形は当業者にとって自明である。選択および記述した実施例は、本開示の原理および実際の適用をより効果的に説明し、かつ当業者に本願を理解させて特定の用途に適する様々な修正付きの様々な実施例を設計するためのものである。



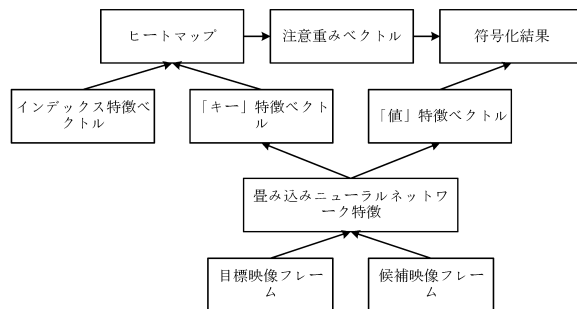
【図 1】



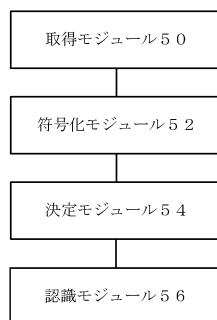
【図 2】



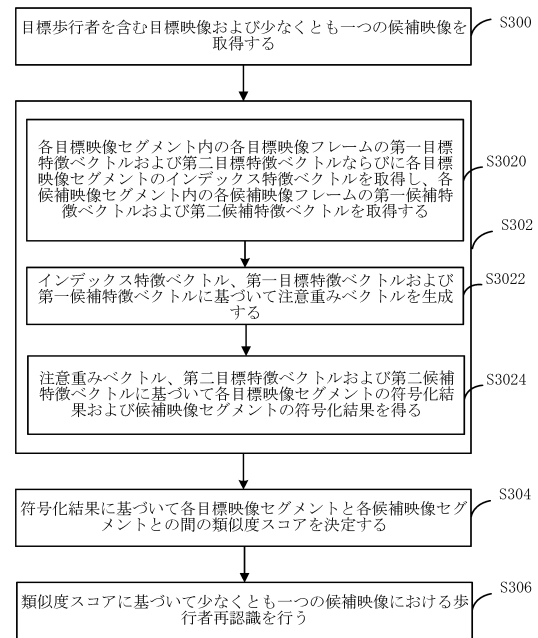
【図 4】



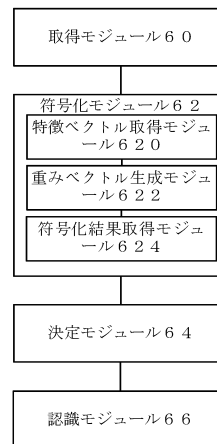
【図 5】



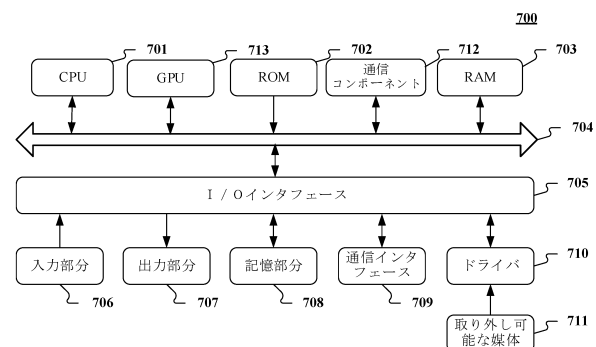
【図 3】



【図 6】



【図 7】



## フロントページの続き

- (72)発明者 チェン ダーボン  
中華人民共和国 ベイジン, ハイディアン ディストリクト, チュンクワンソン イースト ロード, コートヤード 1, ビルディング 3, フロア 7, ルーム 710-712
- (72)発明者 リー ホンション  
中華人民共和国 ベイジン, ハイディアン ディストリクト, チュンクワンソン イースト ロード, コートヤード 1, ビルディング 3, フロア 7, ルーム 710-712
- (72)発明者 シアオ トン  
中華人民共和国 ベイジン, ハイディアン ディストリクト, チュンクワンソン イースト ロード, コートヤード 1, ビルディング 3, フロア 7, ルーム 710-712
- (72)発明者 イー シュワイ  
中華人民共和国 ベイジン, ハイディアン ディストリクト, チュンクワンソン イースト ロード, コートヤード 1, ビルディング 3, フロア 7, ルーム 710-712
- (72)発明者 ワン シャオガン  
中華人民共和国 ベイジン, ハイディアン ディストリクト, チュンクワンソン イースト ロード, コートヤード 1, ビルディング 3, フロア 7, ルーム 710-712

審査官 山田 辰美

- (56)参考文献 特開2017-167970(JP, A)  
特開2017-182437(JP, A)  
Taiqing Wang, Shaogang Gong, Xiatian Zhu, Shengjin Wang, Person Re-identification by Video Ranking, 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV, スイス, springer, 2014年 9月 6日, p.688 - p.703, URL, [https://link.springer.com/content/pdf/10.1007%2F978-3-319-10593-2\\_45.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-319-10593-2_45.pdf)  
Zhedong Zheng, Liang Zheng, Yi Yang, A Discriminatively Learned CNN Embedding for Person Re-identification, JOURNAL OF LATEX CLASS FILES., 米国, 2017年 2月 3日, VOL.14, NO.18, p.1 - p.10, URL, <https://arxiv.org/pdf/1611.05666.pdf>

- (58)調査した分野(Int.Cl., DB名)  
G06T 7/00 - 7/90