

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6873126号

(P6873126)

(45) 発行日 令和3年5月19日(2021.5.19)

(24) 登録日 令和3年4月22日(2021.4.22)

(51) Int.Cl.

F I

G06T 7/246 (2017.01)

G06T 7/246

G06T 7/00 (2017.01)

G06T 7/00 660A

請求項の数 9 (全 23 頁)

(21) 出願番号 特願2018-524732 (P2018-524732)
 (86) (22) 出願日 平成28年11月4日 (2016.11.4)
 (65) 公表番号 特表2018-533805 (P2018-533805A)
 (43) 公表日 平成30年11月15日 (2018.11.15)
 (86) 国際出願番号 PCT/CN2016/104491
 (87) 国際公開番号 W02017/080399
 (87) 国際公開日 平成29年5月18日 (2017.5.18)
 審査請求日 令和1年11月5日 (2019.11.5)
 (31) 優先権主張番号 201510772348.7
 (32) 優先日 平成27年11月12日 (2015.11.12)
 (33) 優先権主張国・地域又は機関
 中国 (CN)

(73) 特許権者 520015461
 アドバンスド ニュー テクノロジーズ
 カンパニー リミテッド
 英国領ケイマン諸島 グランド ケイマン
 ケーワイ1-9008 ジョージ タウ
 ン ホスピタル ロード 27 ケイマン
 コーポレート センター
 (74) 代理人 100188558
 弁理士 飯田 雅人
 (74) 代理人 100205785
 弁理士 ▲高▼橋 史生

最終頁に続く

(54) 【発明の名称】 顔位置追跡方法、装置及び電子デバイス

(57) 【特許請求の範囲】

【請求項1】

顔位置追跡方法であって：

現在の画像フレームにおける第1の顔を検出し、前記現在の画像フレームにおける前記第1の顔の第1の顔領域を取得するステップ(S1)であって、前記現在の画像フレームと次の画像フレームとは連続した画像フレームである、ステップと；

前記第1の顔を含むと予測される前記次の画像フレームにおける予測領域を特定するステップ(S2)であって、前記予測領域は、前記現在の画像フレームにおける前記顔領域のサイズ及び位置を基準とし、前記現在の画像フレームにおける前記顔領域を拡大した長さ及び幅を有する、ステップと；

前記顔領域に対する類似度が所定の要件を満たす初めの事前選択領域を求めて前記予測領域を検索するステップ(S3)であって、2画素以上の大きさの移動する第1のステップサイズに基づいて前記予測領域を検討する、ステップと；

前記初めの事前選択領域を囲む第2のステップサイズの範囲内で、前記顔領域と最も類似性の高い第1の事前選択領域を検索するステップであって、前記第2のステップサイズは前記第1のステップサイズより小さい、ステップと；

前記次の画像フレームに第2の顔領域が存在することを検出するステップ(S4)と；

前記第1の事前選択領域、前記第2の顔領域、及び前記第1の事前選択領域又は前記第2の顔領域を最終的な顔位置追跡結果として用いるかどうかを特定するための所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するステップ(S4)と；

10

20

を備え、

前記第 1 の事前選択領域又は前記第 2 の顔領域を最終的な顔位置追跡結果として用いるかどうかを特定するための所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定する前記ステップは：

前記次の画像フレームにおける前記第 2 の顔領域と前記第 1 の事前選択領域との間の重複係数が 0 である場合、前記第 1 の事前選択領域を前記次の画像フレームの最終の前記顔位置追跡結果として用いるステップと；

前記次の画像フレームにおける前記第 2 の顔領域と前記第 1 の事前選択領域との間の前記重複係数が所定の閾値未満である場合、前記第 1 の事前選択領域を、前記次の画像フレームの最終の前記顔位置追跡結果として用いるステップと；

前記次の画像フレームにおける前記第 2 の顔領域と前記第 1 の事前選択領域との間の前記重複係数が前記所定の閾値以上である場合、前記次の画像フレームにおける前記第 2 の顔領域を前記次の画像フレームの前記顔位置追跡結果として用いるステップと、を備える、

顔位置追跡方法。

【請求項 2】

前記顔領域に対する類似度が前記所定の要件を満たす前記初めの事前選択領域を求めて前記予測領域を検索する前記ステップは：

前記移動する第 1 のステップサイズに基づき、前記予測領域を詳しく検討して前記顔領域の各比較領域を取得するステップと；

前記顔領域と前記各比較領域との類似度を計算するステップと；

前記次の画像フレームにおける前記所定の要件を満たす最大の類似度を有する比較領域を、前記次の画像フレームにおける前記初めの事前選択領域として用いるステップ；を備える、

請求項 1 に記載の顔位置追跡方法。

【請求項 3】

前記顔領域と前記各比較領域との類似度が、現在の比較領域と前記顔領域との間の境界位置の比較、及び指定された経験的閾値に対する前記比較領域と前記次の画像フレームの座標を持つ画像のグレースケール値との間の差に基づいて計算される、

請求項 2 に記載の顔位置追跡方法。

【請求項 4】

現在の画像フレームにおける顔領域を取得する前記ステップが：

少なくとも 2 つの顔が前記現在の画像フレームから検出された場合、前記現在の画像フレームにおける最大領域の顔に対応する領域を、前記現在の画像フレームにおける前記顔領域として選択するステップ；を備える、

請求項 1 に記載の顔位置追跡方法。

【請求項 5】

前記現在の画像フレームにおける前記顔領域の長さ及び幅は、複数画素分拡大される、請求項 1 乃至請求項 4 のいずれか 1 項に記載の顔位置追跡方法。

【請求項 6】

前記現在の画像フレームにおける前記顔領域の長さ及び幅は、複数倍に拡大される、

請求項 1 乃至請求項 4 のいずれか 1 項に記載の顔位置追跡方法。

【請求項 7】

前記現在の画像フレームが複数の顔を含む場合、カメラに最も近い顔が主たる対象である、

請求項 1 乃至請求項 6 のいずれか 1 項に記載の顔位置追跡方法。

【請求項 8】

前記現在の画像フレームにおける前記第 1 の顔の第 1 の顔領域を取得するステップは、分類レベルの量を低減するための AdaBoost 法を用いることで第 1 の顔領域を取得するステップを備える、

10

20

30

40

50

請求項 1 乃至請求項 7 のいずれか 1 項に記載の顔位置追跡方法。

【請求項 9】

請求項 1 乃至請求項 8 のいずれか 1 項に記載の方法を実行するように構成された複数のモジュールを備える、

顔位置追跡のための装置。

【発明の詳細な説明】

【技術分野】

【0001】

本願は、2015年11月12日に提出され「顔位置追跡方法、装置及び電子デバイス」と題された中国特許出願第201510772348.7号の優先権を主張し、上記中国特許出願は参照によってその全体が本願に組み込まれる。

10

【0002】

本願は画像情報データ処理の分野に関し、特に、顔位置追跡方法、装置及び電子デバイスに関する。

【背景技術】

【0003】

顔追跡は通常、ビデオ又は画像シーケンスの中の顔の移動及び大きさの変化を特定する工程である。顔追跡は常に、画像の分析及び認識、画像の監視及び検索、インスタントビデオ通信などにおいて重要な役割を果たし、広く用いられている。

【0004】

20

一般に、顔追跡処理に主として含まれるのは、ビデオ中の顔位置を識別することである。ビデオの録画中、顔が動くと、特定のアルゴリズム、例えば粒子シフト又は平均シフト(Mean shift)、を用いてビデオ中の顔の特定の位置を追跡できる。現在、既存の技術で用いられている顔追跡法における処理に主として含まれるのは、各画像フレーム上での顔検出である、すなわち、各フレームを別々の画像として捉え、各画像フレーム上で顔検出を行って各画像フレームにおける顔位置を計算により取得する。しかし実際の実施では、例えば、ユーザが携帯電話の前向きカメラを使って自撮り画像を撮影する工程において、顔追跡喪失(顔トラッキングロス)又は検出エラーが、通常は、光又はシーンの突然の変化、強い光、計測障害、急速な顔移動などが原因で発生する。その結果、ユーザビデオ監視工程又はビデオ通話工程では、追跡された顔の画像が通常は不連続となり、リアルタイムの滑らかな追跡効果が得られないので、とりわけ処理性能が比較的低い端末デバイスでのユーザエクスペリエンスは大幅に悪化してしまう。当然、既存の技術における顔追跡法は、ユーザの比較的高度な顔追跡要求を満たすことができない。

30

【0005】

既存の技術における顔追跡法では、特に突然の光変化、光干渉、又は急速な顔移動のような複雑なシナリオにおいて顔追跡喪失又はエラーが発生し、それによりビデオの中でぼやけた顔の画像又は不連続な追跡顔画像を生ずる。顔の検出及び追跡並びにユーザエクスペリエンスの有効性が損なわれる可能性がある。

【発明の概要】

【0006】

40

本願は、突然の光変化、光干渉、又は急速な顔移動のような複雑なシナリオにおいても、画像フレーム中の顔領域を正確に突き止め、顔追跡の有効性を高めることができる顔位置追跡の方法、装置及び電子デバイスを提供するためのものである。加えて、顔追跡喪失を回避することで、顔位置追跡効率及びユーザエクスペリエンスを向上させることができるものである。

【0007】

本願において提供する顔位置追跡方法、装置及び電子デバイスは、以下のように実施される。

【0008】

現在の画像フレームにおける顔領域を取得するステップと；

50

前記現在の画像フレームにおける前記顔領域に基づき、前記現在の画像フレームの次の画像フレームにおける顔を含む予測領域を特定するステップと；

前記顔領域に対する類似度が所定の要件を満たす第1の事前選択領域を求めて前記予測領域を検索するステップと；

前記次の画像フレームにおける顔領域を検出し、前記第1の事前選択領域、前記次の画像フレームにおける前記顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するステップと；を含む、顔位置追跡方法が提供される。

【0009】

現在の画像フレームにおける顔領域を検出するように構成された検出モジュールと；

前記現在の画像フレーム内にあり、前記検出モジュールによって検出された前記顔領域に基づき、前記現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を計算するように構成された予測領域計算モジュールと；

前記顔領域に対する類似度が所定の要件を満たす第1の事前選択領域を求めて前記予測領域を検索するように構成された事前選択領域計算モジュールと；

前記第1の事前選択領域、前記検出モジュールによる前記現在の画像フレームの前記次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するように構成された追跡結果選択モジュールと；を含む、顔位置追跡装置が提供される。

【0010】

処理を待つ現在の画像フレームを取得するように構成された情報取得ユニットと；

前記現在の画像フレームにおける顔領域を検出し；前記現在の画像フレームにおける前記検出した顔領域に基づき、前記現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を計算し、前記顔領域に対する類似度が所定の要件を満たす第1の事前選択領域を求めて前記予測領域を検索し；前記第1の事前選択領域、前記処理ユニットによる前記現在の画像フレームの前記次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するように構成された処理ユニットと；

前記処理ユニットが取得した前記顔位置追跡結果を表示するように構成されたディスプレイユニットと；を含む、顔位置追跡電子デバイスが提供される。

【0011】

本願で提供される顔位置追跡方法、装置、及び電子デバイスでは、次の画像フレームにおける顔を含む予測領域の範囲を、現在の画像フレームにおける顔領域に基づいて予測できる。次いで、前の画像フレームにおける顔領域に類似し、所定の要件を満たす（例えば、類似度が最大の）顔領域の事前選択領域を予測領域の範囲内で見つけることができる。そうしたことから、事前に選択された顔に関する情報を、前の画像フレームに基づき、現在のフレームの次の画像フレームから取得できる。更に、本願における解決策では、次の画像フレームにおける顔領域を検出できる。次の画像フレームから顔領域が検出されない場合、前の画像フレームに基づき計算によって取得した事前選択領域を、次の画像フレームにおける顔領域として用いることができる。顔領域が検出できれば、次の画像フレームは少なくとも2つの顔領域を含むため、所定の選択規則に基づいて、次の画像フレームの最終的な顔位置追跡結果として、その要件を満たす顔領域を選択することができる。したがって、本願においては、突然の光変化のような複雑なシナリオのために現在の画像フレームから顔領域が検出されない場合でも、前の画像フレームにおける顔領域に基づき予測により取得した事前選択領域を用いることにより、顔位置を突き止めて追跡できる。そうしたことから、顔検出及び追跡の連続性を確保でき、顔検出及び追跡をより効果的なものにでき、顔追跡のユーザエクスペリエンスを向上させることができる。

【図面の簡単な説明】

【0012】

本願の実施における又は従来技術における技術的解決策をより明確に説明するため、本願の実施又は既存の技術の説明に用いる添付図面を簡単に以下説明する。明らかなことで

10

20

30

40

50

あるが、以下の説明において添付図面は本願の一部の実施を示すに過ぎず、当業者であれば、創造的活動を行うことなくこれらの添付図面から他の図面を導き出し得る。

【0013】

【図1】図1は、本願の実施形態に係る、顔位置追跡方法を示す方法のフローチャートである。

【0014】

【図2】図2は、現在の画像フレームにおける顔領域に基づき、次の画像フレームにおける、顔を含む予測領域を特定することを示す概略図である。

【0015】

【図3】図3は、本願に係る、第1の事前選択領域を求めて予測領域を検索することを示す概略図である。

10

【0016】

【図4】図4は、本願に係る、顔位置追跡結果を特定する選択シナリオを示す概略図である。

【0017】

【図5】図5は、本願の実施形態に係る、第2の事前選択領域を求めて更に検索することを示す概略図である。

【0018】

【図6】図6は、本願の実施形態に係る、顔位置追跡装置のモジュール構造を示す概略図である。

20

【0019】

【図7】図7は、本願の実施形態に係る、事前選択領域計算モジュールのモジュール構造を示す概略図である。

【0020】

【図8】図8は、本願の別の実施形態に係る、事前選択領域計算モジュールのモジュール構造を示す概略図である。

【0021】

【図9】図9は、本願の実施形態に係る、追跡結果選択モジュールのモジュール構造を示す概略図である。

【0022】

30

【図10】図10は、本願の実施形態に係る、顔位置追跡電子装置を示す概略構造図である。

【発明を実施するための形態】

【0023】

当業者が本願における技術的解決策をよりよく理解できるように、本願の実施における添付図面を参照して本願の実施における技術的解決策を明確かつ十分に以下記載する。明らかなことであるが、記載の実施は、本願の実施の全てではなく、一部に過ぎない。本願の実施に基づき当業者が創造的活動なしに獲得する他の全ての実施は、本願の保護範囲に含まれる。

【0024】

40

以下、添付図面を参照して本願における基本的なデータ処理方法を詳細に記載する。図1は、本願の実施形態に係る、顔位置追跡方法を示す方法フローチャートである。本願は、以下の実施又は添付図面において方法の操作ステップを提供するが、この方法は、習慣的又は非創造的な活動に基づく、より多くの又は少数の操作ステップを含むことができる。必要な論理の因果関係が存在しないステップにおいて、これらのステップの実行順序は、本願の実施において提供される実行順序に限定されない。実際の用途において装置又は端末製品によって方法が実行される場合、この方法は、実施若しくは添付図面における方法の順序に基づき実行できる、又は、並行して（例えば、並列プロセッサ又はマルチスレッド処理環境で）実行できる。

【0025】

50

具体的には、図 1 に示すように、本願は、顔位置追跡方法の実施を提供し、この方法は、以下のステップを含むことができる。

【 0 0 2 6 】

S 1 : 現在の画像フレームにおける顔領域を取得する。

【 0 0 2 7 】

一般に、顔位置追跡は、主にカメラ装置、例えば監視デバイス又はユーザの携帯電話のカメラにより録画されたビデオストリーム情報処理で用いられる。本願で提供される顔位置追跡方法は、ビデオストリーム情報処理を含むことができるが、それには限定されない。本願における解決策は、他のアプリケーションシナリオにおける連続画像又は映画フィルムデジタル情報の顔追跡にも用いることができる。本願における解決策を明確に記載するため、この実施は、ユーザが携帯電話の前向きカメラを使用して自撮り画像を撮影するアプリケーションシナリオを用いて説明できる。この実施態様において、まず現在のビデオの現在の画像フレームにおける顔領域が取得される。具体的には、顔検出は、ビデオストリームの最初のビデオ画像フレームで遂行できる。一般に、顔検出は、各画像フレームで顔が検出されるまで遂行する必要がある。一般に、現在の画像フレームから顔が検出されると、画像色パラメータ、画像大きさ、及びフレーム距離などの顔画像に関する情報を取得できる。顔位置は通常、特定の領域で表すことができる。例えば、通常は矩形マス(枠)を用いて顔が検出された領域を表すことができる。

【 0 0 2 8 】

モバイル通信端末又は車両ダッシュボードビデオレコーダのデータ処理能力がパーソナルコンピュータ(P C)のデータ処理能力よりも低い端末製品アプリケーションシナリオにおいて、本願は顔領域を検出する実施態様を提供できる。一の実施態様において、現在の画像フレームにおける顔領域を取得するステップは：

分類レベルの量を低減するための A d a b o o s t 法を用いることにより現在の画像フレームにおける顔領域を検出し取得するステップを含むことができる。

【 0 0 2 9 】

A d a b o o s t は反復アルゴリズムである。A d a b o o s t の一次工程は、同じ学習セット用の異なる分類器(弱い分類器)に学習させるステップと、弱い分類器をより強い最終の分類器(強い分類器)に統合するステップとを含む。A d a b o o s t アルゴリズムは通常、各学習セットにおける各サンプルの分類が正しいか否か、及び過去の全般的な分類の精度に基づき各サンプルの加重値を特定するために用いることができる。次いで、加重値が修正されている新しいデータセットを次層の分類器へ学習のために送信でき、各学習中に取得された分類器は、最終特定用の分類器に統合される。A d a b o o s t カスケード構造における分類器は通常、一連のシリアル分類器で構成される。識別すべきサンプルを特定する場合、前のレベルで分類器により正であると特定されたサンプルのみが更なる処理のために次の分類器に送信される。そうでない場合、負サンプルは直接拒絶される。カスケード構造において、前の分類器は構造がより単純であり、用いる特徴データが少ないが、検出率が高い。対象オブジェクトとは大きく異なる負サンプルは、可能な限りフィルタ除去することができる。後続の分類器は、対象オブジェクトに類似した負サンプルを対象オブジェクトから区別できるようにより多くの特徴データを用い、構造はもっと複雑である。

【 0 0 3 0 】

なお、この実施態様において、A d a b o o s t 法を用いることにより現在の画像フレームにおける顔領域を検出する工程では、顔検出の計算の複雑さを低減し、迅速な顔検出を遂行するため、必要に応じて一定量の分類レベルを適切に低減できる。

【 0 0 3 1 】

本願における顔位置追跡方法の一の実施態様において、ユーザが携帯電話の前向きカメラ又は後向きカメラを用いて能動的に撮影する際に、2人以上の人物が撮影された場合、スクリーン又はディスプレイ上では、撮影された顔のうち主たる顔の比率が大きいことが普通である。この場合、カメラに最も近い顔が主たる対象となり得る。したがって、顔位

10

20

30

40

50

置をより正確に追跡し、顔位置追跡に対するユーザの期待に応えるためには、本願の別の実施態様において、顔位置追跡の際にはカメラに最も近い顔のみを追跡してもよい。一の実施態様の処理プロセスでは、現在の画像フレームにおける最大の顔領域を顔追跡対象として選択できる。したがって、本願の別の実施態様において、現在の画像フレームにおける顔領域を取得するステップは：

現在の画像フレームから少なくとも2つの顔が検出された場合、現在の画像フレームにおける最大領域の顔に対応する領域を現在の画像フレームにおける顔領域として選択するステップを含むことができる。

【0032】

上述の実施態様により、ビデオの現在の画像フレームにおける顔領域を取得できる。

【0033】

S2：現在の画像フレームにおける顔領域に基づき、現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を特定する。

【0034】

現在の画像フレームにおける顔領域を取得した後、現在の画像フレームにおける顔領域に基づき、次の画像フレームにおける、顔を含む予測領域を特定できる。予測領域は、特定のアルゴリズム又は規則に基づき、現在の画像フレームの次の画像フレームにおいて特定された一定範囲の顔領域を含むことができる。図2は、現在の画像フレームにおける顔領域に基づき、次の画像フレームにおける、顔を含む予測領域を特定するステップを示す概略図である。図2に示すように、現在の画像フレームNにおいて、矩形マスAが現在の画像フレームにおける検出された顔領域である。次の画像フレームN+1において、前の画像フレームN（すなわち、現在の画像フレームN）における顔領域の矩形マスAの長さ及び幅を別々にK画素分拡大した後に取得される矩形マスBが、次の画像フレームN+1における、顔を含む予測領域として用いられる。当然、別の実施では、予測領域を、別の方法で特定できる。例えば、現在の画像フレームにおける顔領域の矩形マスAの長さ及び幅を別々に1.5倍に拡大した後に取得される矩形マスを、予測領域として用いることができる。

【0035】

現在の画像フレームの次の画像フレームにおける、顔を含む予測領域は、現在の画像フレームにおける顔領域と特定の規則に基づき特定される。

【0036】

S3：顔領域に対する類似度が所定の要件を満たす第1の事前選択領域を求めて予測領域を検索する。

【0037】

次の画像フレームにおける顔を含み得る予測領域を特定した後、予測領域の範囲内で、顔領域に対する類似度が高い第1の事前選択領域を検索しマッチングすることができる。この顔領域は、前の画像フレーム（すなわち、S2における現在の画像フレームN）から取得できる。

【0038】

この実施態様では、顔領域と次の画像フレームにおける第1の事前選択領域との類似度を計算するためのテンプレートマッチング法を提供できる。一の実施態様の処理プロセスにおいて、現在の画像フレームにおける顔領域は、オリジナルテンプレートとして用いることができ、顔領域の境界を移動窓として設定できる。各窓の移動中に、新しいマッチングテンプレートを取得できる。予測領域の範囲内では、予測領域が1ステップ移動する度に、新しい移動窓領域のデータを取得でき、移動窓領域と顔領域との類似度が計算される。本願において、2つの領域間の類似度を計算又は特定するステップは、特定の方法に限定されず、同一又は類似の機能を実施可能な他の方法を本願に適用することができる。具体的には、現在の画像フレームにおける顔領域に基づき、現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を特定するステップは以下を含むことができる。

【0039】

S 3 0 1 : 第 1 のステップサイズ (刻み幅) に基づき、予測領域を詳しく検討 (トラバース) して顔領域の比較領域を取得する。

【 0 0 4 0 】

第 1 のステップサイズは、実際の顔位置追跡における処理速度又は処理精度要件に基づき設定できる。任意の実施態様においては、データ計算量を減らしてデータ処理速度を向上させることができ、顔位置追跡精度要件を確保できる。この実施態様において、第 1 の移動ステップの値の範囲は 2 画素以上とすることができる。

【 0 0 4 1 】

S 3 0 2 : 顔領域と比較領域との類似度を計算する。

【 0 0 4 2 】

異なるアプリケーションシナリオ又は異なるデータ処理要件に基づき、異なる画像領域間の類似度を計算するために、対応する計算方法及び対応する計算パラメータが選択できる。例えば、顔領域と比較領域との類似度は、異なる領域の画像色、画像質感、又は画像勾配に基づき計算できる。本願は、類似度を計算する実施態様を提供する。具体的には、本願の一実施態様において、 dis で表される顔領域と比較領域との類似度は、下式を用いて計算できる：

【 数 1 】

$$\min X = \max(-left_{ori}, -left_{des})$$

$$\max X = \max(width - left_{ori}, width - left_{des})$$

$$\min Y = \max(-top_{ori}, -top_{des})$$

$$\max Y = \max(height - top_{ori}, height - top_{des})$$

$$sumDis = \left\{ \sum_{i=\max(1, \min X)}^{\min(width, \max X)} \sum_{j=\max(1, \min Y)}^{\min(height, \max X)} \min\{|f(i, j) - g(i, j)|, x\} \right\}$$

$$effectiveNum = [\min(width, \max X) - \max(1, \min X)] * [\min(height, \max X) - \max(1, \min Y)]$$

$$dis = sumDis * (width * height) / effectiveNum$$

【 0 0 4 3 】

上式において、 $left_{ori}$ 、 $left_{des}$ 、 top_{ori} 、及び top_{des} は、それぞれ、顔領域の左側の境界の位置、現在の比較領域の左側の境界の位置、顔領域の上側の境界の位置、及び、現在の比較領域の上側の境界の位置を表す。変数 $width$ は顔領域の幅を表し、 $height$ は顔領域の高さを表し、 $f(i, j)$ は現在の画像フレーム (すなわち、次のフレーム $K + 1$ に先行するフレーム K) の顔領域内の座標が (i, j) である画素のグレースケール値を表し、 $g(i, j)$ は、次の画像フレームの比較領域における座標が (i, j) である画素のグレースケール値を表し、 x は指定された経験的閾値を表し、 dis は顔領域と比較領域との類似度である。上式において、 $\max(a, b)$ は、 a と b のうち大きい方の値が選択されることを表し、 $\min(a, b)$ は、 a と b のうち小さい方の値が選択されることを表す。

【 0 0 4 4 】

S 3 0 3 : 次の画像フレームにおける類似度が所定の要件を満たす比較領域を、次の画像フレームにおける第 1 の事前選択領域として用いる。

【 0 0 4 5 】

この実施態様において、所定の要件を設定し、これを用いて、予測領域において予測要

10

20

30

40

50

件を満たす比較領域を選択できる。例えば、所定の要件を次の通り設定してもよい。すなわち、比較領域と顔領域との類似度が90%を超えている、又は、一定のパーセンテージ以内の比較領域を、類似度の並べ替え後に指定する。例えば、類似度が上から3番目までの比較領域を指定する。本願の任意の実施態様において、類似度が所定の要件を満たす比較領域として挙げられるのは：

次の画像フレームにおける比較領域の中で類似度が最大の比較領域である。

【0046】

図3は、本願による、第1の事前選択領域を求めて予測領域を検索するステップを示す概略図である。図3に示すように、次の画像フレームN+1の予測領域Bにおいて、現在の画像フレームNにおける顔領域Aに対する類似度が要件を満たす第1の事前選択領域Cを見つけることができる。この実施態様において、次のフレームの比較領域の中にあり、現在の画像フレームにおける顔領域に対する類似度が最大である領域のみを、第1の事前選択領域として選択できる。この場合、多数の事前選択領域を選択することに比べ、データ処理量を減らすことができ、事前選択領域をより迅速に選択でき、顔位置処理速度を向上することができる。

【0047】

この実施態様において、テンプレートマッチング法を用いて第1の事前選択領域を計算する際、特定の領域内の画素のグレースケール値に対して加算及び減算をすることができる。別の既存の追跡アルゴリズムと比較して、多量の処理及び格納をする必要がなく、時間計算量及び空間計算量が少ない。この実施は、とりわけ、情報データ処理能力の低い中性能～低性能の携帯電話及び監視デバイスに対して広く適用される。したがって、計算量を効果的に減らすことができ、顔追跡精度を高めることができる。加えて、短距離ビデオ録画の顔追跡環境における、例えば携帯電話の前向きカメラのアプリケーションシナリオにおいて、ユーザが自撮り画像を取る場合、通常、スクリーン内の顔の比率は大きくなる。この実施態様におけるテンプレートマッチング法では、ビデオスクリーンの大きな比率を占める顔領域の効果的な情報を取得でき、他の追跡アルゴリズムと比較して、追跡結果がより信頼性の高いものとなる。

【0048】

上述の実施態様により、顔領域に対する類似度が所定の要件を満たす第1の事前選択領域は、予測領域内で特定の計算方法を用いて検索できる。

【0049】

S4：次の画像フレームにおける顔領域を検出し、第1の事前選択領域、次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき次の画像フレームの顔位置追跡結果を特定する。

【0050】

現在の画像フレームが次の画像フレームに切り替わる際に、次の画像フレームに顔領域が存在するか否かを検出できる。次の画像フレームから顔領域が検出された場合、少なくとも2つの顔領域、すなわち、検出された顔領域と顔追跡予測に基づき取得される第1の事前選択領域とが次の画像フレームから取得される。本願において、次の画像フレームの最終的な顔追跡結果は、この少なくとも2つの顔領域に基づく特定の協調的な計算及び分析によって取得できる。

【0051】

本願において、所定の選択規則を使ってどの顔領域を最終的な顔位置追跡結果として用いるかを特定できる。この実施態様における選択規則としては、次の画像フレームにおける顔領域と第1の事前選択領域のいずれかにおける、次の画像フレームにおける顔領域と第1の事前選択領域との重複領域の比率に応じ、次の画像フレームにおける顔領域又は第1の事前選択領域を選択するための選択規則を挙げることができる。この実施態様において、次の画像フレームの顔領域又は第1の事前選択領域における、顔領域と第1の事前選択領域との間の重複領域の比率を、重複係数Qと定義することができる。

【0052】

図4は、本願による、顔位置追跡結果を特定する選択シナリオを示す概略図である。図4に示すように、矩形マスDは、次の画像フレームにおける検出された顔領域を表すことができ、ここでは検出結果と呼ばれる。矩形マスCは、次の画像フレームにおける第1の事前選択領域を表すことができ、ステップS1～S3又は別の実施態様における追跡計算によって取得されたものであり、ここでは追跡結果と呼ばれる。網掛け領域は、次の画像フレームの最終的な特定された顔位置追跡結果である。検出結果と追跡結果が次の画像フレームに共に存在するとき、検出結果が追跡結果と重複していない、すなわち、重複係数Qが0である場合、図4の4-1に示すように、追跡結果を顔位置追跡結果として用いることができる。検出結果と追跡結果との間に重複領域が存在するが、重複領域が比較的小さく指定された重複要件を満たさない場合、図4の4-2に示すように、追跡結果を顔位置追跡結果として用いることができる。別の場合、検出結果と追跡結果との間に重複領域が存在し、重複領域が比較的大きく指定された重複要件を満たす、例えば、領域の95%が重複している場合、図4の4-3に示すように、検出結果を顔位置追跡結果として選択できる。当然、次の画像フレームから顔領域が検出されない場合は、図4の4-4に示すように、追跡結果を直接顔位置追跡結果として用いることができる。

10

【0053】

したがって、任意の実施態様において、次の画像フレームにおける顔領域検出と所定の選択規則とに基づき、次の画像フレームの顔位置追跡結果を特定するステップは：

次の画像フレームにおける顔領域の検出結果として、顔領域が検出されない場合、第1の事前選択領域を次の画像フレームの顔位置追跡結果として用いるステップ；

20

次の画像フレームにおける検出された顔領域と第1の事前選択領域との間の重複係数が0である場合、第1の事前選択領域を次の画像フレームの顔位置追跡結果として用いるステップ；

次の画像フレームにおける検出された顔領域と第1の事前選択領域との間の重複係数が所定の閾値未満である場合、第1の事前選択領域を次の画像フレームの顔位置追跡結果として用いるステップ；又は

次の画像フレームにおける検出された顔領域と第1の事前選択領域との間の重複係数が所定の閾値以上である場合、次の画像フレームにおける検出された顔領域を次の画像フレームの顔位置追跡結果として用いるステップ：のいずれかを含むことができる。

【0054】

30

この実施は、検出結果及び追跡結果から最終の顔位置追跡結果を選択するための実施方法を提供する。本願のこの実施態様においては、急速な顔移動、突然の光変化、又は強い光干渉等の複雑な環境で、顔位置を正確かつ迅速に追跡できる。本実施態様では、ある画像フレームにおいてフレーム喪失が発生し顔が検出されない場合でも顔位置を追跡し特定することが可能であり、その結果、顔位置の連続的追跡効果を実施でき、顔追跡を円滑にできる。フレーム喪失が発生しない場合であっても、顔追跡効果の有効性及びユーザエクスペリエンスを向上させるために、所定の選択規則に基づき、画像フレームにおける検出結果と追跡結果から、より適切な領域を顔位置追跡結果として選択できる。

【0055】

なお、本願における現在のフレーム、次のフレーム、前のフレーム、及び前のフレームに先行するフレームは、実際の用途における画像フレーム情報処理対象を説明するための相対的概念であるとみなすことができる。ビデオストリームにおけるある瞬間の画像フレームを現在の画像フレーム(N)として注目できる場合、対応する次のフレームを(N+1)番目の画像フレームとしてもよく、前のフレームを(N-1)番目の画像フレームとしてもよい。(N+1)番目の画像フレームに対する追跡が完了した後、(N+2)番目の画像フレームにおける顔位置を更に追跡し処理することができる。この場合、現在の画像フレームが(N+1)番目の画像フレームであり、それに対応して、現在の画像フレーム(N+1)の次の画像フレームを(N+2)番目の画像フレームとしてもよい。

40

【0056】

連続する画像フレームの処理中、現在の画像フレームにおける顔位置に対する追跡処理

50

が完了した後、現在の画像フレームの処理結果を次の画像フレームにおける顔位置追跡のための参照情報又は初期化情報として用いることにより、次の画像フレームにおける顔位置を更に追跡できる。アプリケーションシナリオによっては、ビデオストリーム中の1秒当たりの処理が必要な画像フレームの量が、通常で10フレームを超えることもあり、場合によっては数十フレームとなる。顔位置の追跡中、突然の光変化又は急速な顔移動に起因するフレームNでの顔追跡喪失により顔が検出されない場合、前のフレームN-1の検出又は処理結果から取得した顔領域を、顔追跡喪失が発生したフレームNから追跡された顔領域として用いることができる。関連する実施態様としては、さらに、前のフレームN-1からも顔が検出又は追跡されない場合、フレームN-2を検出又は処理して取得した顔領域結果を用いることもでき、以下同様に処理する。当然、所定の特定規則に基づき、多数の連続したフレームにおいて顔が検出されない場合、ビデオの録画範囲に顔が捉えられていないと特定できる。

10

【0057】

上記の実施態様では、次の画像フレームにおいて顔を含む予測領域を検索する場合、必要に応じて第1のステップサイズの値を決定できる。例えば、予測領域を毎回2画素分又は5画素分移動させることができる。一般に、ステップサイズが大きいほど前の画像フレームの顔領域に類似した領域を検索する速度が速いことを意味し、処理に必要なデータ量は少ない。ステップサイズが小さいほど検索精度が高いことを意味する。第1の移動ステップの値の範囲が2画素以上である実施態様においては、第1の事前選択領域を検索する精度を更に向上させるため、本願において提供する別の好適な実施において、方法は、以下を更に含むことができる：

20

【0058】

S304：顔領域に対する類似度が最大の第2の事前選択領域を求めて、第1の事前選択領域を囲む第2のステップサイズの範囲内を検索する。ここで第2のステップサイズは第1のステップサイズ未満である。この場合、正確な検索によって取得される第2の事前選択領域を、次の画像フレームの顔位置領域追跡結果として用いることができる。検出結果と第2の事前選択領域とのいずれかをこの後に最終の顔位置追跡結果と特定してもよい。

【0059】

これに対応して、第1の事前選択領域、次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、次の画像フレームの顔位置追跡結果を特定するステップは：第2の事前選択領域、次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、次の画像フレームの顔位置追跡結果を特定するステップを含む。

30

【0060】

特定の用途において、例えば、第1のステップ（刻み量）は2画素であり、類似度が最大の領域を取得するため、第1の事前選択領域を囲む1画素の範囲内における顔領域と前の画像フレームにおける顔領域との類似度をこの実施態様において計算してもよい。具体的には、この実施態様において、第2の事前選択領域は、ステップS302における類似度計算法を用いて計算できる。当然、2つの領域間の類似度を特定するための他の計算方法が除外されることはない。これらの計算方法の詳細については、ここでは記載しない。図5は、本願の実施形態に係る、第2の事前選択領域を求めて検索するステップを示す概略図である。図5に示すように、矩形マスCは、予測領域のステップサイズとして2画素を用いることにより特定される顔領域の第1の事前選択領域であり、矩形マスDは、第1の事前選択領域Cの1画素分右上の比較領域C_{ru}である。当然、第1の事前選択領域を囲む1画素の範囲は、第1の事前選択領域を1画素分下に移動させることによって形成される比較領域C_d、第1の事前選択領域を1画素分上に移動させることによって形成される比較領域C_u、第1の事前選択領域を1画素分左下角に向けて移動させることによって形成される比較領域C_{ld}などを含むことができる。第1の事前選択領域を囲む1画素の比較領域と顔領域との類似度を計算することができ、類似度が最大の比較領域を第2の事前選択領域として選択できる。

40

50

【 0 0 6 1 】

この実施態様において、第 1 の事前選択領域は、画像比較及び検索における計算量を効果的に減らして顔位置追跡のデータ処理速度を高めるため、比較的大きな第 1 のステップサイズを設定することによって計算される。この実施態様において、第 1 の事前選択領域の結果に基づき、第 1 のステップサイズよりも小さい第 2 のステップサイズを用いることにより、より正確な検索を遂行してより正確な追跡結果を有する第 2 の事前選択領域を取得できる。そうしたことから、迅速な検索を実施でき、顔追跡の精度を向上させ、それによって顔追跡の有効性を向上することができる。

【 0 0 6 2 】

本願における顔位置追跡方法に基づき、本願は、顔位置追跡装置を提供する。図 6 は、本願の実施形態に係る、顔位置追跡装置のモジュール構造を示す概略図である。図 6 に示すように、この装置は、

10

現在の画像フレームにおける顔領域を検出するように構成された検出モジュール 1 0 1 と；

現在の画像フレーム内にあり、検出モジュール 1 0 1 によって検出された顔領域に基づき、現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を計算するように構成された予測領域計算モジュール 1 0 2 と；

顔領域に対する類似度が所定の要件を満たす第 1 の事前選択領域を求めて予測領域を検索するように構成された事前選択領域計算モジュール 1 0 3 と；

第 1 の事前選択領域、検出モジュール 1 0 1 による現在の画像フレームの次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、次の画像フレームの顔位置追跡結果を特定するように構成された追跡結果選択モジュール 1 0 4 と；を含むことができる。

20

【 0 0 6 3 】

本願における顔位置追跡装置の実施態様において、検出モジュール 1 0 1 は、カメラ装置によって取得した画像フレームにおける顔領域を長時間連続して検出できる。例えば、ビデオストリームにおいて 1 秒当たり 1 5 フレームのビデオ画像が撮影され、顔位置の追跡中に画像の現在のフレーム（N 番目のフレーム）における顔領域を検出できる。画像の現在のフレーム（N 番目のフレーム）の情報データに対する検出及び追跡処理が完了した後、画像の次のフレーム（（N + 1）番目のフレーム）における顔領域を更に検出できる。

30

【 0 0 6 4 】

本願における装置の検出モジュール 1 0 1 の一の実施態様において、現在の画像フレームにおける顔領域を、分類レベルの量を減らすための A d a b o o s t 法を用いることにより検出し取得できる。そうしたことから、顔検出中のデータ計算量を減らすことができ、顔位置追跡の位置特定及び処理の速度を高めることができる。

【 0 0 6 5 】

本願における装置の別の実施態様において、顔位置追跡処理の際には、カメラに最も近い顔のみを追跡できる。一の実施態様の処理プロセスにおいて、現在の画像フレームにおける最大の顔領域を顔追跡対象として選択できる。したがって、本願における装置の別の実施態様において、検出モジュール 1 0 1 が現在の画像フレームにおける顔領域を検出するステップは：

40

現在の画像フレームから少なくとも 2 つの顔が検出された場合、現在の画像フレームにおける最大領域の顔に対応する領域を現在の画像フレームにおける顔領域として選択するステップを含む。

【 0 0 6 6 】

本願は、計算に基づく事前選択領域計算モジュール 1 0 3 による第 1 の事前選択領域を検索する方法に限定されない。図 7 は、本願の実施形態に係る、装置の事前選択領域計算モジュール 1 0 3 のモジュール構造を示す概略図である。図 7 に示すように、事前選択領域計算モジュール 1 0 3 は、比較領域モジュール 1 0 3 1、類似度計算モジュール 1 0 3

50

2、及び第1の事前選択モジュール1033を含むことができる。

【0067】

比較領域モジュール1031は、指定された第1のステップに基づき、予測領域を詳しく検討して顔領域の比較領域を取得するように構成されている。

【0068】

類似度計算モジュール1032は、顔領域と比較領域との類似度を計算するように構成されている。

【0069】

本願における装置の一の実施において、類似度計算モジュール1032は、下式を用いることにより顔領域と比較領域との類似度 dis を計算できる：

10

【数1】

$$\min X = \max(-left_{ori}, -left_{des})$$

$$\max X = \max(width - left_{ori}, width - left_{des})$$

$$\min Y = \max(-top_{ori}, -top_{des})$$

$$\max Y = \max(height - top_{ori}, height - top_{des})$$

$$sumDis = \left\{ \sum_{i=\max(1, \min X)}^{\min(width, \max X)} \sum_{j=\max(1, \min Y)}^{\min(height, \max X)} \min\{|f(i, j) - g(i, j)|, x\} \right\}$$

20

$$effectiveNum = [\min(width, \max X) - \max(1, \min X)] * [\min(height, \max X) - \max(1, \min Y)]$$

$$dis = sumDis * (width * height) / effectiveNum$$

【0070】

上式において $left_{ori}$ 、 $left_{des}$ 、 top_{ori} 、及び top_{des} は、それぞれ顔領域の左側の境界の位置、現在の比較領域の左側の境界の位置、顔領域の上側の境界の位置、及び現在の比較領域の上側の境界の位置を表す。変数 $width$ は顔領域の幅を表し、 $height$ は顔領域の高さを表し、 $f(i, j)$ は、現在の画像フレームにおける顔領域内の座標が (i, j) である画素のグレースケール値を表し、 $g(i, j)$ は、次の画像フレームの比較領域における座標が (i, j) である画素のグレースケール値を表し、 x は、指定された経験的閾値を表し、 dis は顔領域と比較領域との類似度である。

30

【0071】

第1の事前選択モジュール1033は、次の画像フレームにおける類似度が所定の要件を満たす比較領域を次の画像フレームにおける第1の事前選択領域として用いるように構成されている。

40

【0072】

装置に含まれる類似度計算モジュール1032が用いる式は、コンピュータ読み取り可能なプログラム言語、例えばC言語を用いることにより、装置/モジュール上での特定の実施工程において実施できる、又は、必要に応じてある種のハードウェア構造を用いるハードウェア及びソフトウェアの形態で実施できる。

【0073】

任意の実施態様において、第1の事前選択モジュール1033によって指定される所定の要件を、顔領域に対する最大の類似度を用いて設定してもよい。したがって、別の実施態様において、第1の事前選択モジュール1033における類似度が所定の要件を満たす

50

比較領域として挙げられるのは：

次の画像フレームにおける比較領域の中で類似度が最大の比較領域である。

【 0 0 7 4 】

上記実施態様において、比較領域モジュール 1 0 3 1 によって指定される第 1 のステップは、本願における顔位置追跡装置の処理速度又は処理精度についての要件に応じて設定できる。本願における装置の一の実施態様において、第 1 のステップサイズの値の範囲は、2 画素以上となるように設定できる。

【 0 0 7 5 】

本願は、顔位置追跡装置の好適な実施態様を更に提供できる。図 8 は、本願の別の実施形態に係る、事前選択領域計算モジュールのモジュール構造を示す概略図である。図 8 に示すように、事前選択領域計算モジュール 1 0 3 は：

顔領域に対する類似度が最大の第 2 の事前選択領域を求めて、第 1 の事前選択領域を囲む第 2 のステップサイズであって、第 1 のステップサイズ未満である第 2 のステップサイズの範囲内を検索するように構成された第 2 の事前選択モジュール 1 0 3 4 を更に含むことができる。

【 0 0 7 6 】

これに対応して、追跡結果選択モジュール 1 0 4 が、第 1 の事前選択領域、検出モジュール 1 0 1 による現在の画像フレームの次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき次の画像フレームの顔位置追跡結果を特定することは、追跡結果選択モジュール 1 0 4 が、第 2 の事前選択領域、検出モジュール 1 0 1 による現在の画像フレームの次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、次の画像フレームの顔位置追跡結果を特定することを含む。

【 0 0 7 7 】

この実施態様における顔位置追跡装置において、第 1 の事前選択領域の結果に基づき、第 1 のステップサイズ未満の第 2 のステップサイズの範囲内で、より正確な検索をすることができ、追跡結果がより正確な第 2 の事前選択領域が取得される。そうしたことから、迅速な検索処理を実施することができ、顔位置追跡の精度を高め、それによって顔追跡の有効性を高めることができる。

【 0 0 7 8 】

図 9 は、本願の実施形態に係る、追跡結果選択モジュール 1 0 4 のモジュール構造を示す概略図である。図 9 に示すように、追跡結果選択モジュール 1 0 4 は、検出・計算モジュール 1 0 4 1 及び選択モジュール 1 0 4 2 を含むことができる。

【 0 0 7 9 】

検出・計算モジュール 1 0 4 1 は、次の画像フレームにおける顔領域を検出し、次の画像フレームにおける顔領域が検出された場合、次の画像フレームの顔領域と第 1 の事前選択領域との重複係数 Q を計算するように構成されている。この実施態様において、重複係数 Q は、顔領域と第 1 の事前選択領域のいずれかにおける、次の画像フレームにおける顔領域と第 1 の事前選択領域との重複領域の比率で表せる。

【 0 0 8 0 】

選択モジュール 1 0 4 2 は、検出・計算モジュール 1 0 4 1 により次の画像フレームから顔領域が検出されない、検出・計算モジュール 1 0 4 1 が計算した重複係数が 0 である、あるいは検出・計算モジュール 1 0 4 1 が計算した重複係数が所定の閾値未満である、という条件のうちの少なくとも 1 つが満たされる場合、第 1 の事前選択領域を、次の画像フレームの顔位置追跡結果として用いるように構成されている。あるいは、選択モジュール 1 0 4 2 は、検出・計算モジュール 1 0 4 1 が計算した重複係数が所定の閾値以上である場合、次の画像フレーム内にあり、検出モジュール 1 0 1 が検出した顔領域を次の画像フレームの顔位置追跡結果として用いるように構成されている。

【 0 0 8 1 】

この実施態様は、検出結果及び追跡結果から最終の顔位置追跡結果を選択することに対する解決策を提供する。本願のこの実施態様において、急速な顔移動、突然の光変化、又

10

20

30

40

50

は強い光干渉のような複雑な環境で、顔位置を正確かつ迅速に追跡できる。この実施態様において、ある画像フレームにおいてフレーム喪失が発生して顔が検出されない場合でも顔位置を追跡し特定することができ、その結果、顔位置の連続的追跡効果を実施して円滑な顔追跡を確保できる。フレーム喪失が発生しない場合でも、顔追跡の有効性及びユーザエクスペリエンスを向上させるために、所定の選択規則に基づき、画像フレームにおける検出結果と追跡結果からより適切な領域を顔位置追跡結果として選択できる。

【0082】

本願における顔位置追跡方法又は装置は、多数の端末デバイスに適用してより迅速で、正確、かつ円滑な顔位置追跡を遂行できる。例えば、そうしたデバイスとしては、ビデオカメラデバイス、監視デバイス、及びAndroid（登録商標）システム又はiOSシステムに基づくモバイル通信端末の連続画像フレーム用顔位置追跡デバイスを挙げることができる。したがって、本願は、顔位置追跡電子デバイスを更に提供し、電子デバイスは、ビデオ画像フレームを検出し取得するためのカメラ装置、ビデオ再生用ディスプレイ、情報データ処理用処理ユニットなどを含むことができる。具体的には、図10は、本願の実施形態に係る、顔位置追跡電子装置を示す概略構造図である。図10に示すように、電子デバイスは：

処理を待つ現在の画像フレームを取得するように構成された情報取得ユニット1と；

現在の画像フレームにおける顔領域を検出し；現在の画像フレームにおける検出した顔領域に基づき、現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を計算し、顔領域に対する類似度が所定の要件を満たす第1の事前選択領域を求めて予測領域を検索し；第1の事前選択領域、処理ユニットによる現在の画像フレームの次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、次の画像フレームにおける顔位置追跡結果を特定するように構成された処理ユニット2と；

処理ユニット2が取得した顔位置追跡結果を表示するように構成されたディスプレイユニット3と；を含むことができる。

【0083】

本願における電子デバイスの情報取得ユニット1は、モバイル端末の前向きカメラ、後向きカメラ、又は監視カメラ装置を含むことができる。別のアプリケーションシナリオにおいては、リアルタイムで取得された又は事前に取得された画像情報データをコンピュータが処理する実施が更に含まれ、例えば、コンピュータが、ビデオ情報に対して顔位置追跡処理を遂行する。処理ユニット2は、中央処理ユニット（CPU）を含むことができ、当然、論理処理能力、論理ゲート回路、集積回路などを有する単一チップマイクロコンピュータを更に含むことができる。ディスプレイユニット3は一般に、ディスプレイ、モバイル端末ディスプレイスクリーン、投影デバイスなどを含むことができる。

【0084】

顔検出方法、モジュールユニット間のデータ交換などのデータ処理、及び情報表示が本願の内容において記載されるが、本願は、一般論又は実施で述べたデータ処理及び情報表示に限定されない。本願の実施における上記の記載は、本願の一部の実施の応用に過ぎず、何らかの基準及び方法に基づきわずかに修正された処理方法も、本願の実施における解決策を実施するために用いることができる。当然、本願の実施における処理方法のステップに従う他の非創造的変更も、同じ用途を実施するために用いることができる。ここではその詳細については触れない。

【0085】

本願は、実施態様又はフローチャートにおける方法の操作ステップを提供するが、この方法は、習慣的又は非創造的活動に基づき、それよりも多数又は少数の操作ステップを含むことができる。実施態様において列挙したステップの順序は、ステップの実行順序の一種に過ぎず、一意的な実行順序を表すものではない。実際の用途において装置又はクライアント製品により方法が実行される場合、方法は、実施態様又は添付図面における方法の順序に基づき実行でき、あるいは、並行して（例えば、並行プロセッサ又はマルチスレッド処理環境）実行できる。

【 0 0 8 6 】

上記の実施において記載したユニット又はモジュールは、具体的にはコンピュータチップ又はエンティティにより実施することができ、あるいは、ある種の機能を有する製品により実施することができる。説明を容易にするため、上記の装置及びデバイスは、様々なモジュール及び様々なユニットを用いることにより説明されている。当然、本願の実施に際し、多数のモジュールの機能を、1つ以上のソフトウェア及び/またはハードウェア、例えば、第1の事前選択モジュール及び第2の事前選択モジュールにおいて実施でき、あるいは、同じ機能を実施するモジュールを、多数のサブモジュール又はサブユニットの組合せを用いて実施できる。

【 0 0 8 7 】

10

当業者にとっても周知であるが、コンピュータ読み取り可能なプログラムコードを用いることによりコントローラを実装することに加えて、方法ステップに対して論理プログラミングを遂行してコントローラが論理ゲート、スイッチ、専用集積回路、プログラマブル論理コントローラ、及び埋込型マイクロコントローラの形態で同じ機能を実施できる。したがって、コントローラは、ハードウェアコンポーネントとみなすことができ、コントローラに含まれ様々な機能を実施するために用いられる装置は更に、ハードウェアコンポーネントにおける構造体とみなすことができる。あるいは、様々な機能を実施するための装置は、場合により方法を実施するためのソフトウェアモジュールとハードウェアコンポーネントにおける構造体の両方とみなすこともできる。

【 0 0 8 8 】

20

本願は、C言語、又はAndroid（登録商標）デザインプラットフォーム若しくはiOSデザインプラットフォームに基づくプログラムモジュールのようなコンピュータが実行するコンピュータ実行可能な指令の一般的文脈において記載することができる。一般に、プログラムモジュールは、特定の作業を実行する、あるいは特定の抽象データ型を実施するルーチン、プログラム、オブジェクト、コンポーネント、データ構造、型などを含む。本願は更に、通信ネットワークを用いることにより接続された遠隔処理デバイスによって作業が実行される分散型コンピューティング環境に適用することができる。分散型コンピューティング環境において、プログラムモジュールは、格納デバイスを含むローカル及びリモートコンピュータ格納媒体に配置することができる。

【 0 0 8 9 】

30

上記実施の記載から分かることであるが、本願は、ユニバーサルハードウェアプラットフォームに加えて、ソフトウェアを用いて実施できることを当業者は明確に理解できる。こうした理解に基づき、本願における技術的解決策は本質的に、あるいは従来技術に貢献する部分は、ソフトウェア製品の形態で実施できる。ソフトウェア製品は、ROM/RAM、磁気ディスク、又は光ディスクなどの記録媒体に格納することができ、本願の実施形態又は実施形態の一部に記載した方法を遂行するようにコンピュータデバイス（パーソナルコンピュータ、モバイル端末、サーバ、又はネットワークデバイスとすることができる）に命令する幾つかの指令を含む。

【 0 0 9 0 】

40

本明細書における実施は全て、進捗的な方法で記載されている。実施における同一又は類似の部分については、これらの実施を参照することができ、各実施形態は、他の実施形態との違いに焦点を当てている。本願は、多くの汎用又は専用コンピュータシステム環境若しくは構成、例えば、パーソナルコンピュータ、サーバコンピュータ、手持ち式デバイス若しくはポータブルデバイス、タブレットデバイス、モバイル通信端末、マルチプロセッサシステム、マイクロプロセッサシステム、プログラム可能電子デバイス、ネットワークPC、小型コンピュータ、メインフレームコンピュータ、及び上記システム又はデバイスのいずれかを含む分散型コンピューティング環境に適用することができる。

【 0 0 9 1 】

実施の形態を用いて本願を記載したが、本願は本願の精神から逸脱することなく多くの変形及び変更を有し、添付の請求項は、本願の精神から逸脱することなくこれらの変形及

50

び変更を含むことは、当業者には周知である。

以下、本発明の実施の態様の例を列挙する。

[第 1 の局面]

顔位置追跡方法であって：

現在の画像フレームにおける顔領域を取得するステップと；

前記現在の画像フレームにおける前記顔領域に基づき、前記現在の画像フレームの次の画像フレームにおける顔を含む予測領域を特定するステップと；

前記顔領域に対する類似度が所定の要件を満たす第 1 の事前選択領域を求めて前記予測領域を検索するステップと；

前記次の画像フレームにおける顔領域を検出し、前記第 1 の事前選択領域、前記次の画像フレームにおける前記顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するステップと；を備える、

顔位置追跡方法。

[第 2 の局面]

前記顔領域に対する類似度が所定の要件を満たす第 1 の事前選択領域を求めて前記予測領域を検索する前記ステップが：

第 1 の移動ステップに基づき、前記予測領域を詳しく検討して前記顔領域の比較領域を取得するステップと；

前記顔領域と前記比較領域との類似度を計算するステップと；

前記次の画像フレームにおける類似度が前記所定の要件を満たす比較領域を、前記次の画像フレームにおける前記第 1 の事前選択領域として用いるステップ；を備える、

第 1 の局面に記載の顔位置追跡方法。

[第 3 の局面]

前記顔領域と前記比較領域との類似度が、下式を用いて計算される：

【 数 1 】

$$\min X = \max(-left_{ori}, -left_{des})$$

$$\max X = \max(width - left_{ori}, width - left_{des})$$

$$\min Y = \max(-top_{ori}, -top_{des})$$

$$\max Y = \max(height - top_{ori}, height - top_{des})$$

$$sumDis = \left\{ \sum_{i=\max(1, \min X)}^{\min(width, \max X)} \sum_{j=\max(1, \min Y)}^{\min(height, \max X)} \min\{|f(i, j) - g(i, j)|, x\} \right\}$$

$$effectiveNum = [\min(width, \max X) - \max(1, \min X)] * [\min(height, \max X) - \max(1, \min Y)]$$

$$dis = sumDis * (width * height) / effectiveNum$$

(ここで、 $left_{ori}$ 、 $left_{des}$ 、 top_{ori} 、及び top_{des} は、それぞれ、前記顔領域の左側の境界の位置、現在の比較領域の左側の境界の位置、前記顔領域の上側の境界の位置、及び前記現在の比較領域の上側の境界の位置を表し； $width$ は、前記顔領域の幅を表し、 $height$ は、前記顔領域の高さを表し、 $f(i, j)$ は、前記現在の画像フレームにおける前記顔領域における座標が (i, j) である画素のグレースケール値を表し、 $g(i, j)$ は、前記次の画像フレームの比較領域における座標が (i, j) である画素のグレースケール値を表し； x は、指定された経験的閾値を表し、 dis は、前記顔領域と前記比較領域との類似度である)、

第 2 の局面に記載の顔位置追跡方法。

[第 4 の局面]

類似度が前記所定の要件を満たす前記比較領域は、前記次の画像フレームにおける比較領域の中で類似度が最大の比較領域を含む、

第 2 の局面に記載の顔位置追跡方法。

[第 5 の局面]

前記第 1 の移動ステップの値の範囲が 2 画素以上である、

第 2 の局面に記載の顔位置追跡方法。

[第 6 の局面]

顔領域に対する類似度が最大の第 2 の事前選択領域を求めて、第 1 の事前選択領域を囲む第 2 のステップサイズの範囲内を検索するステップであって、前記第 2 のステップサイズは第 1 のステップサイズ未満である、顔領域に対する類似度が最大の第 2 の事前選択領域を求めて、第 1 の事前選択領域を囲む第 2 のステップサイズの範囲内を検索するステップと；

これに対応して、前記第 1 の事前選択領域、前記次の画像フレームにおける前記顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定する前記ステップが、前記第 2 の事前選択領域、前記次の画像フレームにおける前記顔領域の前記検出結果、及び前記所定の選択規則に基づき、前記次の画像フレームの前記顔位置追跡結果を特定するステップ；を更に備える、

第 2 乃至 5 の局面のいずれかに記載の顔位置追跡方法。

[第 7 の局面]

前記第 1 の事前選択領域、前記次の画像フレームにおける前記顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定する前記ステップが；

前記次の画像フレームにおける前記顔領域の前記検出結果として、顔領域が検出されない場合、前記第 1 の事前選択領域を前記次の画像フレームの前記顔位置追跡結果として用いるステップと；

前記次の画像フレームにおける前記検出された顔領域と前記第 1 の事前選択領域との間の重複係数が 0 である場合、前記第 1 の事前選択領域を前記次の画像フレームの前記顔位置追跡結果として用いるステップと；

前記次の画像フレームにおける前記検出された顔領域と前記第 1 の事前選択領域との間の前記重複係数が所定の閾値未満である場合、前記第 1 の事前選択領域を、前記次の画像フレームの前記顔位置追跡結果として用いるステップと；

前記次の画像フレームにおける前記検出された顔領域と前記第 1 の事前選択領域との間の前記重複係数が前記所定の閾値以上である場合、前記次の画像フレームにおける前記検出された顔領域を前記次の画像フレームの前記顔位置追跡結果として用いるステップと；の何れかを備える、

第 1 の局面に記載の顔位置追跡方法。

[第 8 の局面]

現在の画像フレームにおける顔領域を取得する前記ステップが；

少なくとも 2 つの顔が前記現在の画像フレームから検出された場合、前記現在の画像フレームにおける最大領域の顔に対応する領域を、前記現在の画像フレームにおける前記顔領域として選択するステップ；を備える、

第 1 の局面に記載の顔位置追跡方法。

[第 9 の局面]

顔位置追跡装置であって；

現在の画像フレームにおける顔領域を検出するように構成された検出モジュールと；

前記現在の画像フレーム内にあり、前記検出モジュールによって検出された前記顔領域に基づき、前記現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を計算するように構成された予測領域計算モジュールと；

前記顔領域に対する類似度が所定の要件を満たす第 1 の事前選択領域を求めて前記予測領域を検索するように構成された事前選択領域計算モジュールと；

前記第 1 の事前選択領域、前記検出モジュールによる前記現在の画像フレームの前記次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するように構成された追跡結果選択モジュールと；を備える、

顔位置追跡装置。

[第 10 の局面]

前記事前選択領域計算モジュールが：

指定された第 1 のステップに基づき、前記予測領域を詳しく検討して前記顔領域の比較領域を取得するように構成された比較領域モジュールと；

前記顔領域と前記比較領域との類似度を計算するように構成された類似度計算モジュールと；

前記次の画像フレームにおける類似度が前記所定の要件を満たす比較領域を、前記次の画像フレームにおける前記第 1 の事前選択領域として用いるように構成された第 1 の事前選択モジュールと；を備える、

第 9 の局面に記載の顔位置追跡装置。

[第 11 の局面]

前記顔領域と前記比較領域との類似度が、下式を用いて計算される：

【数 1】

$$\min X = \max(-left_{ori}, -left_{des})$$

$$\max X = \max(width - left_{ori}, width - left_{des})$$

$$\min Y = \max(-top_{ori}, -top_{des})$$

$$\max Y = \max(height - top_{ori}, height - top_{des})$$

$$sumDis = \left\{ \sum_{i=\max(1, \min X)}^{\min(width, \max X)} \sum_{j=\max(1, \min Y)}^{\min(height, \max X)} \min\{|f(i, j) - g(i, j)|, x\} \right\}$$

$$effectiveNum = [\min(width, \max X) - \max(1, \min X)] * [\min(height, \max X) - \max(1, \min Y)]$$

$$dis = sumDis * (width * height) / effectiveNum$$

(ここで、 $left_{ori}$ 、 $left_{des}$ 、 top_{ori} 、及び top_{des} は、それぞれ前記顔領域の左側の境界の位置、現在の比較領域の左側の境界の位置、前記顔領域の上側の境界の位置、及び前記現在の比較領域の上側の境界の位置を表し； $width$ は、前記顔領域の幅を表し、 $height$ は、前記顔領域の高さを表し、 $f(i, j)$ は、前記現在の画像フレームにおける前記顔領域における座標が (i, j) である画素のグレースケール値を表し、 $g(i, j)$ は、前記次の画像フレームの比較領域における座標が (i, j) である画素のグレースケール値を表し； x は、指定された経験的閾値を表し、 dis は、前記顔領域と前記比較領域との類似度である)、

第 10 の局面に記載の顔位置追跡装置。

[第 12 の局面]

前記第 1 の事前選択モジュールにおける類似度が前記所定の要件を満たす前記比較領域は、前記次の画像フレームにおける比較領域の中で類似度が最大の比較領域を含む、

第 10 の局面に記載の顔位置追跡装置。

[第 1 3 の局面]

前記第 1 の移動ステップの値の範囲は 2 画素以上である、
第 1 0 の局面に記載の顔位置追跡装置。

[第 1 4 の局面]

前記顔領域に対する類似度が最大の第 2 の事前選択領域を求めて、前記第 1 の事前選択領域を囲む第 2 のステップサイズの範囲内を検索するように構成された追跡結果選択モジュールであって、前記第 2 のステップサイズは前記第 1 のステップサイズ未満である、追跡結果選択モジュールと；

これに対応して、前記追跡結果選択モジュールが、前記第 1 の事前選択領域、前記検出モジュールによる前記現在の画像フレームの前記次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定することは、前記追跡結果選択モジュールが、前記第 2 の事前選択領域、前記検出モジュールによる前記現在の画像フレームの前記次の画像フレームにおける前記顔領域前記検出結果、及び前記所定の選択規則に基づき、前記次の画像フレームの前記顔位置追跡結果を特定することを含む、

第 1 0 乃至 1 3 の局面のいずれかに記載の顔位置追跡装置。

[第 1 5 の局面]

前記追跡結果選択モジュールが：

前記次の画像フレームにおける前記顔領域を検出し、前記次の画像フレームにおける前記顔領域が検出された場合、前記次のフレームの顔領域と前記第 1 の事前選択領域との重複係数を計算するように構成された検出・計算モジュールと；

前記検出・計算モジュールにより前記次の画像フレームから顔領域が検出されない、前記検出・計算モジュールが計算した前記重複係数が 0 である、あるいは前記検出・計算モジュールが計算した前記重複係数が所定の閾値未満である、という条件のうちの少なくとも 1 つが満たされる場合、前記第 1 の事前選択領域を前記次の画像フレームの前記顔位置追跡結果として用い；前記検出・計算モジュールが計算した前記重複係数が、前記所定の閾値以上である場合、前記次の画像フレーム内にあり、前記検出モジュールが検出した前記顔領域を前記次の画像フレームの前記顔位置追跡結果として用いるように構成された選択モジュールと；を備える、

第 9 の局面に記載の顔位置追跡装置。

[第 1 6 の局面]

前記検出モジュールが現在の画像フレームにおける顔領域を検出することが：

少なくとも 2 つの顔が前記現在の画像フレームから検出された場合、前記現在の画像フレームにおける最大領域の顔に対応する領域を前記現在の画像フレームにおける前記顔領域として選択するステップを備える、

第 9 の局面に記載の顔位置追跡装置。

[第 1 7 の局面]

顔位置追跡電子デバイスであって：

処理を待つ現在の画像フレームを取得するように構成された情報取得ユニットと；

前記現在の画像フレームにおける顔領域を検出し；前記現在の画像フレームにおける前記検出した顔領域に基づき、前記現在の画像フレームの次の画像フレームにおける、顔を含む予測領域を計算し、前記顔領域に対する類似度が所定の要件を満たす第 1 の事前選択領域を求めて前記予測領域を検索し；前記第 1 の事前選択領域、前記処理ユニットによる前記現在の画像フレームの前記次の画像フレームにおける顔領域検出結果、及び所定の選択規則に基づき、前記次の画像フレームの顔位置追跡結果を特定するように構成された処理ユニットと；

前記処理ユニットが取得した前記顔位置追跡結果を表示するように構成されたディスプレイユニットと；を備える、

顔位置追跡電子デバイス。

10

20

30

40

50

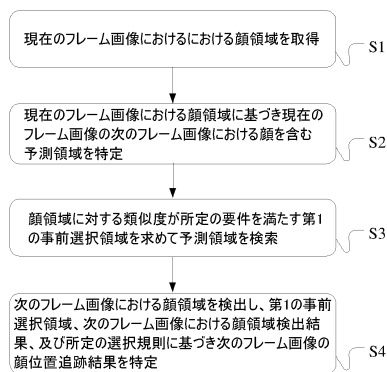
【符号の説明】

【0092】

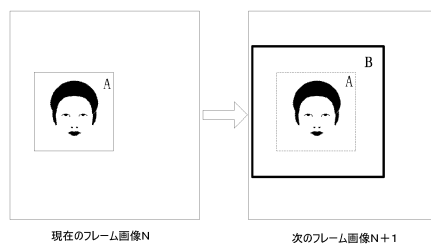
- 1 情報取得ユニット
- 2 処理ユニット
- 3 ディスプレイユニット
- 101 検出モジュール
- 102 予測領域計算モジュール
- 103 事前選択領域計算モジュール
- 104 追跡結果選択モジュール
- 1031 比較領域モジュール
- 1032 類似度計算モジュール
- 1033 第1の事前選択モジュール
- 1034 第2の事前選択モジュール
- 1041 検出・計算モジュール
- 1042 選択モジュール

10

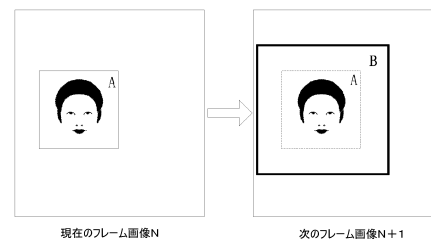
【図1】



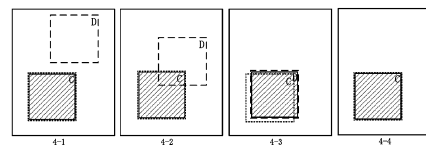
【図2】



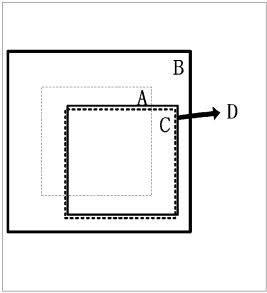
【図3】



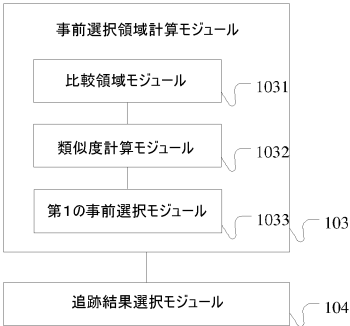
【図4】



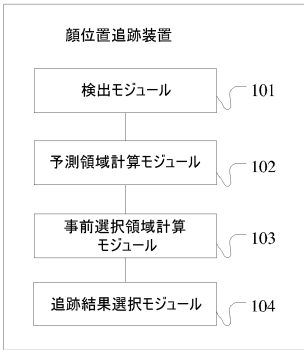
【図 5】



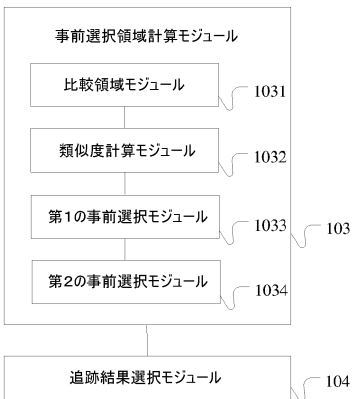
【図 7】



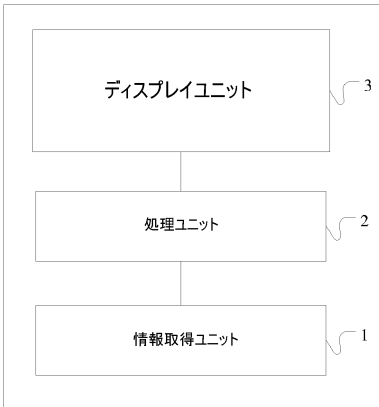
【図 6】



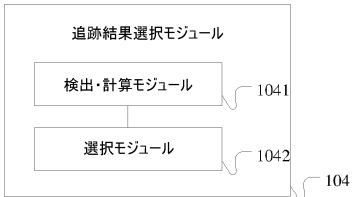
【図 8】



【図 10】



【図 9】



フロントページの続き

(72)発明者 ワン, ナン

中華人民共和国 310099, ハンチョウ, ナンバー18 ワンタン ロード, ファンロン
タイムズ プラザ, ビルディング ビー 17エフ, アンツ パテント チーム内

(72)発明者 ドゥ, チージュン

中華人民共和国 310099, ハンチョウ, ナンバー18 ワンタン ロード, ファンロン
タイムズ プラザ, ビルディング ビー 17エフ, アンツ パテント チーム内

(72)発明者 チャン, ユー

中華人民共和国 310099, ハンチョウ, ナンバー18 ワンタン ロード, ファンロン
タイムズ プラザ, ビルディング ビー 17エフ, アンツ パテント チーム内

審査官 真木 健彦

(56)参考文献 特開2008-217714(JP, A)

特開2010-191784(JP, A)

米国特許出願公開第2014/0286527(US, A1)

特開2011-170711(JP, A)

特開2010-021943(JP, A)

特開2009-104284(JP, A)

特開2015-114917(JP, A)

特開2013-210824(JP, A)

特開2010-141616(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/246

G06T 7/00

G06T 1/00