

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7635002号
(P7635002)

(45)発行日 令和7年2月25日(2025.2.25)

(24)登録日 令和7年2月14日(2025.2.14)

(51)国際特許分類

F I

G 0 6 T 7/20 (2017.01) G 0 6 T 7/20 3 0 0 Z

H 0 4 N 7/18 (2006.01) H 0 4 N 7/18 G

請求項の数 7 (全24頁)

(21)出願番号	特願2021-8803(P2021-8803)	(73)特許権者	000004352
(22)出願日	令和3年1月22日(2021.1.22)		日本放送協会
(65)公開番号	特開2022-112831(P2022-112831 A)		東京都渋谷区神南2丁目2番1号
(43)公開日	令和4年8月3日(2022.8.3)	(73)特許権者	399060908
審査請求日	令和5年12月22日(2023.12.22)		一般財団法人NHK財団
			東京都世田谷区用賀四丁目10番1号
		(74)代理人	100141139
			弁理士 及川 周
		(74)代理人	100171446
			弁理士 高田 尚幸
		(74)代理人	100114937
			弁理士 松本 裕幸
		(74)代理人	100171930
			弁理士 木下 郁一郎
		(72)発明者	望月 貴裕

最終頁に続く

(54)【発明の名称】 顔追跡装置およびプログラム

(57)【特許請求の範囲】

【請求項1】

映像に含まれる顔領域の動き予測の処理を行うことによって、前記顔領域の動き予測結果の情報を出力する映像解析処理部と、

前記映像に含まれる前記顔領域の特徴量を抽出するとともに、前記特徴量に基づいて前記顔領域の顔認識処理を行うことによって、前記顔領域の位置情報である第1位置情報と、顔認識処理の結果として人物を識別する情報である人物識別情報と前記人物識別情報に対応する認識スコアとの組である第1の組と、を出力する顔認識処理部と、

前記顔領域の時間方向のトラックを求める追跡処理部と、
を備え、

前記トラックは、前記顔領域の位置情報である第2位置情報と、前記顔領域の人物識別情報と認識スコアとの組である第2の組と、を持ち、

前記追跡処理部は、前記映像解析処理部から渡される前記顔領域の動き予測結果の情報に基づいて前記顔領域が動いた後の前記第2位置情報を求め、前記顔認識処理部から渡される前記第1の組と、前記トラックが持つ前記第2の組と、に基づいて、且つ前記第1位置情報と前記第2位置情報との関係にも基づいて、前記トラックを更新することによって、前記トラックに、新たな前記第2位置情報と新たな前記第2の組と、を付与するものであり、

前記映像解析処理部は、さらに、前記顔領域がボケ領域であるか否かの判定を行い、

前記追跡処理部は、前記ボケ領域である前記顔領域に関しては、前記顔領域のトラック

を更新する際には前記認識スコアの値を最低値として、

前記顔認識処理部は、前記顔領域がボケ領域である場合の前記特徴量と、当該ボケ領域の前記特徴量に関連付けられた特殊な前記人物識別情報と、を関連付けた情報にアクセス可能と予めしておき、

前記映像解析処理部は、前記顔認識処理部が前記顔領域についての顔認識処理を行った結果として前記ボケ領域の前記特徴量に関連付けられた特殊な前記人物識別情報を同定した場合に、当該顔領域が前記ボケ領域であると判定する、

顔追跡装置。

【請求項 2】

前記顔追跡装置は、前記映像解析処理部による処理を、前記顔認識処理部による処理よりも高頻度で行い、

10

前記映像解析処理部は、複数回の前記動き予測の処理の結果である前記動き予測結果の情報を蓄積し、

前記追跡処理部は、前記映像解析処理部から渡される蓄積された前記動き予測結果の情報に基づいて処理を行う、

請求項 1 に記載の顔追跡装置。

【請求項 3】

前記映像解析処理部は、さらに、前記映像内のシーンチェンジの検出を行い、

前記追跡処理部は、前記シーンチェンジが検出された場合には、前記トラックを初期状態にリセットする、

20

請求項 1 または 2 に記載の顔追跡装置。

【請求項 4】

前記追跡処理部は、前記第 1 位置情報と前記第 2 位置情報との重なり度合いに基づいて、前記トラックを更新する、

請求項 1 から 3 までのいずれか一項に記載の顔追跡装置。

【請求項 5】

前記トラックの情報に前記人物識別情報と前記認識スコアとの組の情報を付与した形の、前記トラックの集合のデータを整形して出力する出力整形処理部、

をさらに備える請求項 1 から 4 までのいずれか一項に記載の顔追跡装置。

【請求項 6】

30

前記人物識別情報と、前記人物識別情報に関連付けられた人物名表記文字列と、前記人物識別情報に関連付けられたサムネイル画像の指定と、のいずれかの検索キーを取得し、取得した前記検索キーに基づいて、前記出力整形処理部が出力した前記トラックの集合のデータを検索することによって、特定の人物識別情報に関連付けられた前記トラックを特定する、検索処理部、

をさらに備える請求項 5 に記載の顔追跡装置。

【請求項 7】

請求項 1 から 6 までのいずれか一項に記載の顔追跡装置、としてコンピューターを機能させるためのプログラム。

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、顔追跡装置およびプログラムに関する。

【背景技術】

【0002】

画像中の顔を認識して人物を特定する顔認識技術は、近年の AI（人工知能）技術の進展によって高精度化している。顔認識技術は、セキュリティ分野では実用化の域に達している。

【0003】

特許文献 1 は、顔認識技術について記載している。特許文献 1 によれば、顔認識技術は

50

、学習用の顔画像を収集し、人物の違いを認識するための顔特徴量を画像に基づいて計算し、その顔特徴量の類似度に基づいて顔の認識を行うというのが基本である。

【0004】

上記のような画像での顔認識技術を映像に適用することによって、人物識別情報などのメタデータを効率的に映像に付与する取り組みが、放送事業者等を中心に広がりを見せている。画像処理を時間的なつながりの有る映像処理に拡張する際に、有益な処理が物体領域の追跡処理である。

【0005】

特許文献2は、追跡処理の結果を顔認識に利用する手順を記載している。

【0006】

特許文献3は、予め準備した画像とのテンプレートマッチングによる処理を利用した追跡処理を採用している。

【0007】

対象を人の顔に特定しない、映像中の一般的な物体追跡技術についても、これまでに種々の提案が行われてきている。一般的な追跡処理の代表例は、特許文献4や特許文献5に記載されている。

【0008】

特許文献4には、異なるフレームで検出した物体領域の特徴量の類似度に応じて、領域の対応関係を明らかにして、物体の追跡を行う処理が記載されている。

【0009】

特許文献5では、パーティクルフィルタと呼ばれる統計的な処理を行う手法が記載されている。特許文献5に記載の追跡処理の手法は、パーティクルフィルタによって、領域間の対応関係に重み付けを行うことによって追跡の信頼度を向上させようとしている。

【0010】

近年では、追跡処理技術の分野でも、深層学習処理を活用した処理の導入が図られている。深層学習処理の活用により、追跡精度が大きく向上している。

【0011】

非特許文献1は、深層学習処理を用いた追跡処理の手法を多く記載している。また、非特許文献1においても引用されている非特許文献2もまた、深層学習処理を用いた追跡処理について記載している。

【先行技術文献】

【特許文献】

【0012】

【文献】特開2017-033372号公報

【文献】特開2014-191479号公報

【文献】特開2014-021896号公報

【文献】国際公開WO2018/163243 A1

【文献】特開2019-153112号公報

【非特許文献】

【0013】

【文献】Gioele Ciaparrone, Francisco Luque Sanchez, Siham Tabik, Luigi Troiano, Roberto Tagliaferri, Francisco Herrera, Deep Learning in video multi-object tracking: A survey, ArXiv:1907.12740, 2019年, [令和3年(西暦2021年)1月1日検索], インターネット<URL: <https://arxiv.org/pdf/1907.12740.pdf>>

【文献】Nicolai Wojke, Alex Bewley, Dietrich Paulus, Simple Online and Real time Tracking with a Deep Association Metric, arXiv:1703.07402, 2017年, [令和3年(西暦2021年)1月1日検索], インターネット<URL: <https://arxiv.org/pdf/1703.07402.pdf>>

【発明の概要】

【発明が解決しようとする課題】**【 0 0 1 4 】**

しかしながら、それぞれの背景技術には課題が存在する。

【 0 0 1 5 】

特許文献 2 や特許文献 3 に記載されている追跡処理は、顔認識処理の付随的な処理にとどまっており、認識結果と追跡とのセマンティックな結合はない。特許文献 2 や特許文献 3 の技術では、顔認識処理の出力を基に映像中の人物特定を簡単に行えるようにはなっていないという問題がある。

【 0 0 1 6 】

非特許文献 1 や非特許文献 2 に記載された追跡技術は、あくまで同一物体領域のトラック情報を求めるものである。非特許文献 1 や非特許文献 2 では、顔認識などの認識結果と追跡結果とでの、高度なセマンティック情報の融合処理については記載されていない。

【 0 0 1 7 】

また、従来技術では、映像におけるボケ領域や、シーンチェンジに対応することはできないという問題もある。本発明は、映像内の人物が出現している区間で、一貫した顔認識を実現しようとするものである。

【 0 0 1 8 】

本発明は、上記のような事情を考慮して為されたものであり、追跡処理結果を顔認識処理結果と融合することにより、トラック情報に人物名などのセマンティック性の高い情報を与えることのできる顔追跡装置およびプログラムを提供しようとするものである。

【課題を解決するための手段】**【 0 0 1 9 】**

[1] 上記の課題を解決するため、本発明の一態様による顔追跡装置は、映像に含まれる顔領域の動き予測の処理を行うことによって、前記顔領域の動き予測結果の情報を出力する映像解析処理部と、前記映像に含まれる前記顔領域の特徴量を抽出するとともに、前記特徴量に基づいて前記顔領域の顔認識処理を行うことによって、前記顔領域の位置情報である第 1 位置情報と、顔認識処理の結果として人物を識別する情報である人物識別情報と前記人物識別情報に対応する認識スコアとの組である第 1 の組と、を出力する顔認識処理部と、前記顔領域の時間方向のトラックを求める追跡処理部と、を備え、前記トラックは、前記顔領域の位置情報である第 2 位置情報と、前記顔領域の人物識別情報と認識スコアとの組である第 2 の組と、を持ち、前記追跡処理部は、前記映像解析処理部から渡される前記顔領域の動き予測結果の情報に基づいて前記顔領域が動いた後の前記第 2 位置情報を求め、前記顔認識処理部から渡される前記第 1 の組と、前記トラックが持つ前記第 2 の組と、に基づいて、且つ前記第 1 位置情報と前記第 2 位置情報との関係にも基づいて、前記トラックを更新することによって、前記トラックに、新たな前記第 2 位置情報と新たな前記第 2 の組と、を付与する、ものである。

【 0 0 2 0 】

[2] また、本発明の一態様は、上記の顔追跡装置において、前記顔追跡装置は、前記映像解析処理部による処理を、前記顔認識処理部による処理よりも高頻度で行い、前記映像解析処理部は、複数回の前記動き予測の処理の結果である前記動き予測結果の情報を蓄積し、前記追跡処理部は、前記映像解析処理部から渡される蓄積された前記動き予測結果の情報に基づいて処理を行う、というものである。

【 0 0 2 1 】

[3] また、本発明の一態様は、上記の顔追跡装置において、前記映像解析処理部は、さらに、前記映像内のシーンチェンジの検出を行い、前記追跡処理部は、前記シーンチェンジが検出された場合には、前記トラックを初期状態にリセットする、ものである。

【 0 0 2 2 】

[4] また、本発明の一態様は、上記の顔追跡装置において、前記映像解析処理部は、さらに、前記顔領域がボケ領域であるか否かの判定を行い、前記追跡処理部は、前記ボケ領域である前記顔領域に関しては、前記顔領域のトラックを更新する際には前記認識スコ

10

20

30

40

50

アの値を最低値（最悪スコア値）とする、ものである。

【 0 0 2 3 】

〔 5 〕また、本発明の一態様は、上記の顔追跡装置において、前記顔認識処理部は、前記顔領域がボケ領域である場合の前記特徴量と、当該ボケ領域の前記特徴量に関連付けられた特殊な前記人物識別情報と、を関連付けた情報にアクセス可能と予めしておき、前記映像解析処理部は、前記顔認識処理部が前記顔領域についての顔認識処理を行った結果として前記ボケ領域の前記特徴量に関連付けられた特殊な前記人物識別情報を同定した場合に、当該顔領域が前記ボケ領域であると判定する、というものである。

【 0 0 2 4 】

〔 6 〕また、本発明の一態様は、上記の顔追跡装置において、前記追跡処理部は、前記第 1 位置情報と前記第 2 位置情報との重なり度合いに基づいて、前記トラックを更新する、というものである。

10

【 0 0 2 5 】

〔 7 〕また、本発明の一態様は、上記の顔追跡装置において、前記トラックの情報に前記人物識別情報と前記認識スコアとの組の情報を付与した形の、前記トラックの集合のデータを整形して出力する出力整形処理部、をさらに備えるものである。

【 0 0 2 6 】

〔 8 〕また、本発明の一態様は、上記〔 7 〕の顔追跡装置において、前記人物識別情報と、前記人物識別情報に関連付けられた人物名表記文字列と、前記人物識別情報に関連付けられたサムネイル画像の指定と、のいずれかの検索キーを取得し、取得した前記検索キーに基づいて、前記出力整形処理部が出力した前記トラックの集合のデータを検索することによって、特定の人物識別情報に関連付けられた前記トラックを特定する、検索処理部、をさらに備えるものである。

20

【 0 0 2 7 】

〔 9 〕また、本発明の一態様は、映像に含まれる顔領域の動き予測の処理を行うことによって、前記顔領域の動き予測結果の情報を出力する映像解析処理部と、前記映像に含まれる前記顔領域の特徴量を抽出するとともに、前記特徴量に基づいて前記顔領域の顔認識処理を行うことによって、前記顔領域の位置情報である第 1 位置情報と、顔認識処理の結果として人物を識別する情報である人物識別情報と前記人物識別情報に対応する認識スコアとの組である第 1 の組と、を出力する顔認識処理部と、前記顔領域の時間方向のトラックを求める追跡処理部と、を備え、前記トラックは、前記顔領域の位置情報である第 2 位置情報と、前記顔領域の人物識別情報と認識スコアとの組である第 2 の組と、を持ち、前記追跡処理部は、前記映像解析処理部から渡される前記顔領域の動き予測結果の情報に基づいて前記顔領域が動いた後の前記第 2 位置情報を求め、前記顔認識処理部から渡される前記第 1 の組と、前記トラックが持つ前記第 2 の組と、に基づいて、且つ前記第 1 位置情報と前記第 2 位置情報との関係にも基づいて、前記トラックを更新することによって、前記トラックに、新たな前記第 2 位置情報と新たな前記第 2 の組と、を付与する、顔追跡装置、としてコンピューターを機能させるプログラムである。

30

【発明の効果】

【 0 0 2 8 】

本発明によれば、顔領域の画像としての特徴量の情報と、位置や動き等も考慮したトラック（複数フレームにまたがる顔領域の追跡）の情報との、両方に基づいて、新たなフレームにおいて、トラックを同定したり、顔認識を行ったりすることができる。これにより、追跡の精度の向上と、顔認識の精度の向上の、両方が期待できる。

40

【図面の簡単な説明】

【 0 0 2 9 】

【図 1】本発明の一実施形態による顔追跡装置の概略機能構成を示すブロック図である。

【図 2】同実施形態による顔追跡装置における、映像解析処理と追跡処理との時間的關係を示す概略図である。

【図 3】同実施形態において、顔認識処理と一体化したボケ判定処理のための、顔データ

50

ベースの構築方法を模式的に示す概略図である。

【図４】同実施形態における追跡処理部による追跡処理の手順を示すフローチャートである。

【図５】同実施形態における、あるトラック（顔領域のトラック）についての認識スコアの時間推移の例を示す概略図である。

【図６】同実施形態の顔追跡装置による追跡処理の結果として得られるデータ構成の一例を示す概略図である。

【図７】同実施形態による人物名あるいはサムネイル画像を起点として、当該人物が登場する映像コンテンツのシーンを取り出すための情報経路の例を示す概略図である。

【図８】同実施形態における、上記（図７）の映像シーン検索のためのＧＵＩ（グラフィカルユーザーインターフェース）の一例を示す概略図である。

【図９】同実施形態による顔追跡装置の内部構成の例を示すブロック図である。

【発明を実施するための形態】

【００３０】

次に、本発明の一実施形態について、図面を参照しながら説明する。本実施形態の顔追跡装置は、時間的に連続して出現する人物の顔認識、オクルージョンや一時的な人物の消失にも対応した頑健な顔領域追跡手法、映像中のシーンチェンジや動きボケなどの変化に対応した頑健な手法といった技術要素を実現する。これにより、顔認識の結果に基づいて、トラック情報に人物名などのセマンティック性の高い情報を与えることを可能とする。また、本実施形態の顔追跡装置は、映像を対象とする人物検索を容易にするための、認識処理結果のデータを出力する。これにより、顔認識結果を活用したアプリケーションソフトウェアの開発が容易になる。本実施形態の顔追跡装置は、そのために、追跡処理と顔認識処理を統合したデータ構造で、処理結果を出力する。

【００３１】

本実施形態は、時間的に連続する画像（言い換えれば、時間的に前後するフレーム画像）に基づいて、時間的な関係性を考慮して、顔を追跡する。また、顔追跡処理と顔認識処理とを統合する。これにより、本実施形態の顔追跡装置は、顔追跡に関する情報（顔領域の位置の情報や、顔領域の動き予測に関する情報）を、顔認識処理の精度向上のために利用する。また、本実施形態の顔追跡装置は、顔認識処理に関する情報（顔領域の特徴量の類似性（特徴量間の距離））を、顔追跡（トラッキング）の精度向上のために利用する。

【００３２】

図１は、本実施形態による顔追跡装置の概略機能構成を示すブロック図である。図示するように、顔追跡装置１は、映像取得部１０と、映像解析処理部２０と、顔データベース３０と、顔認識処理部４０と、追跡処理部５０と、出力整形処理部６０と、を含んで構成される。これらの各機能部は、例えば、コンピューターと、プログラムとで実現することが可能である。また、各機能部は、必要に応じて、記憶手段を有する。記憶手段は、例えば、プログラム上の変数や、プログラムの実行によりアロケーションされるメモリーである。また、必要に応じて、磁気ハードディスク装置やソリッドステートドライブ（ＳＳＤ）といった不揮発性の記憶手段を用いるようにしてもよい。また、各機能部の少なくとも一部の機能を、プログラムではなく専用の電子回路として実現してもよい。

【００３３】

顔追跡装置１は、映像を入力し、映像内に含まれる顔を認識し、顔認識結果と映像情報とを統合したデータを出力する。顔追跡装置１が出力するデータは、所定のデータ構造で構成されるものである。顔追跡装置１を構成する各部の機能は、次に説明する通りである。

【００３４】

映像取得部１０は、処理対象（顔追跡対象）の映像を外部から取得する。ここで「外部」とは、他の装置等であってもよいし、データを記録する記録媒体等であってもよい。なお、映像は、所定のフレームレート（例えば、３０フレーム毎秒、あるいは６０フレーム毎秒等）で連続するフレーム画像のシーケンスである。

【００３５】

10

20

30

40

50

映像解析処理部 20 は、映像取得部 10 が取得した映像の解析を行う。具体的には、映像解析処理部 20 は、映像を構成するフレーム画像間での物体や人物等の動き予測処理（動きパラメータ値の抽出）を行ったり、シーンチェンジ解析処理を行ったり、フレーム画像内の領域ごとのボケ検出を行ったりする。映像解析処理部 20 は、映像解析処理の結果を追跡処理部 50 に供給する。

【0036】

言い換えれば、映像解析処理部 20 は、映像に含まれる顔領域の動き予測の処理を行うことによって、前記顔領域の動き予測結果の情報（動き予測のモデルの情報）を出力する。

【0037】

顔データベース 30 は、人物を識別するための情報である人物 ID と、顔画像の特徴量と、を関連付けて記憶するデータベースである。データベース自体は既存の技術を用いて実現される。データベースは、記憶手段として、例えば磁気ディスク装置や半導体記憶装置などを用いることができる。なお、顔データベース 30 に登録される特徴量は、収集された大量の顔画像を基に、機械学習処理等によって得られる情報である。言い換えれば、顔データベース 30 は、顔認識モデルの情報を記憶する。

10

【0038】

なお、顔データベース 30 が、人物 ID の他に、当該人物の属性情報（「政治家」、「俳優」などといったジャンル等の情報）を合わせて保持するようにしてもよい。

【0039】

顔認識処理部 40 は、映像取得部 10 が取得した映像中の顔領域の部分を検出し、その顔領域部分から計算される特徴量と、顔データベース 30 が記憶している特徴量とを照合することによって顔認識の結果を出力する。顔認識処理部 40 は、顔認識処理の結果として、人物を特定する情報である人物 ID を出力する。あるいは、顔認識処理部 40 は、映像内の特定の顔領域に対応して、人物 ID ごとのスコアを出力するようにしてもよい。このスコアは、当該顔領域に含まれる顔が、その人物 ID によって識別される人の顔であることの尤度を表す。なお、顔認識処理自体は、既存技術を用いて行うことができる。

20

【0040】

顔認識処理部 40 は、言い換えれば、映像に含まれる顔領域の特徴量を抽出するとともに、その特徴量に基づいて顔領域の顔認識処理を行うことによって、顔領域の位置情報（第 1 位置情報と呼ぶ）と、顔認識処理の結果として人物を識別する情報である人物識別情報と人物識別情報に対応する認識スコアとの組（第 1 の組と呼ぶ）と、を出力する。

30

【0041】

追跡処理部 50 は、映像解析処理部 20 からの映像解析処理結果を活用して、顔認識処理部 40 が出力するフレーム画像内の顔領域の追跡処理を行う。これにより、追跡処理部 50 は、時間的に連続する顔領域トラックを求める。ここで「時間的に連続する」とは、隣接し合うフレーム画像間で連続することを意味する。これにより、顔認識の精度が向上する。

【0042】

つまり、追跡処理部 50 は、顔領域の時間方向のトラックを求める。ここで、トラックは、追跡情報（軌跡情報）のデータである。トラックは、顔領域の位置情報（第 2 位置情報と呼ぶ）と、顔領域の人物識別情報と認識スコアとの組（第 2 の組と呼ぶ）と、を持つ。追跡処理部 50 は、映像解析処理部 20 から渡される顔領域の動き予測結果の情報に基づいて顔領域が動いた後の第 2 位置情報を求め、顔認識処理部から渡される第 1 の組と、元々トラックが持っている第 2 の組と、に基づいて、且つ前記第 1 位置情報と前記第 2 位置情報との関係（重なり度合い）にも基づいて、トラックを更新する。これによって、追跡処理部は、トラックに、新たな第 2 位置情報と新たな第 2 の組（人物識別情報と認識スコアの組）とを付与する。

40

【0043】

出力整形処理部 60 は、所定のデータ構造に整形した形で、追跡結果データを出力する。出力整形処理部 60 が出力するデータの構造は、映像検索処理等の映像インタラクシ

50

ンシステムから利用しやすい構造であり、顔認識結果の情報と、映像に関する情報とを含むものである。出力整形処理部 60 が出力するデータ構造により、出力データを利用する機能（アプリケーションソフトウェア等）は、映像中の人物をキーとした検索などの操作を効果的に実現する。

【0044】

つまり、出力整形処理部 60 は、追跡処理部 50 から渡される追跡処理の結果のデータを整形し、出力する。一例として、出力整形処理部 60 は、トラックの情報に人物識別情報（人物 ID）と認識スコアとの組の情報を付与した形の、トラックの集合のデータを整形して出力する。人物識別情報と認識スコアの組は、複数であってよい。

【0045】

顔追跡装置 1 は、映像解析処理部 20 や、追跡処理部 50 や、出力整形処理部 60 の処理によって、従来技術では為し得なかった効果を生じさせる。次に、映像解析処理部 20、追跡処理部 50、出力整形処理部 60 のそれぞれの、より詳細な処理内容について説明する。

【0046】

映像解析処理部 20 は、映像のシーンチェンジの検出や、フレーム画像間におけるグローバルな動きおよび領域ごとのローカルな動きの予測や、領域がボケているかどうかを判定する処理を行う。これらの映像解析処理の結果は、顔認識処理部 40 の結果とともに追跡処理部 50 および出力整形処理部 60 に供給され、顔認識精度向上のために活用される。

【0047】

映像解析処理部 20 が行うシーンチェンジの検出は、映像内に映っている内容が大きく変わる切れ目を検出する処理である。シーンチェンジが生じるときには、追跡処理部 50 が行う映像内の顔の追跡処理をリセットすることが必要となる。シーンチェンジの検出のために、映像解析処理部 20 は、前後するフレーム画像間の相関度合いを計算する。シーンチェンジが生じていない状況においては、前後するフレーム画像間の相関度合いは所定の閾値以上のレベルを維持する。シーンチェンジが生じるときには、前後するフレーム画像間の相関度合いが上記閾値を大きく下回る。映像解析処理部 20 は、上記の相関度合いがこの閾値を下回っているか否かに基づいて、シーンチェンジが生じたか否かを判定する。このシーンチェンジの検出自体は、既存技術を利用して行うことができる。時刻（ $t-1$ ）のフレーム画像と時刻 t のフレーム画像との間（ここでの t は整数値）でシーンチェンジが生じたか否かを表す情報を、映像解析処理部 20 は出力する。映像解析処理部 20 は、例えば変数 `is_sc_change` の値を出力することにより、シーンチェンジの検出結果を追跡処理部 50 に伝える。`is_sc_change = 1` は「シーンチェンジ有り」を表す。`is_sc_change = 0` は「シーンチェンジ無し」を表す。

【0048】

映像解析処理部 20 が行う動き予測の処理は、前後のフレーム画像間で画像内に映っている物や人の動きの程度を予測するものである。映像解析処理部 20 は、動きの程度の予測値を、追跡処理部 50 に渡す。これにより、追跡処理部 50 は、映像解析処理部 20 から渡された予測値にしたがった動き補償処理を行い、追跡処理の精度を向上させることができる。

【0049】

なお、映像解析処理部 20 は、動き予測の処理として、画面全体の動きであるグローバル動き（撮影用カメラのパン等）と、検出された顔領域に特化した動きであるローカル動きとの、2種類の動きを予測する。映像解析処理部 20 は、グローバル動きの予測として、フレーム画像間の相関度に基づいて、画面全体のグローバル動きをモデル化したパラメータを計算する。グローバル動きのモデルの代表的な例として、アフィンモデルがある。アフィンモデル自体は、既存技術によって処理可能なモデルである。グローバル動きについてのパラメータは、`globalM` というマトリックスで表現される。このマトリックス `globalM` を用いた動き補償は、下の式（1）を用いて計算可能である。

【0050】

10

20

30

40

50

$$p(t) = \text{globalM} \cdot p(t-1) \cdots (1)$$

【0051】

この式(1)において、 $p(t-1)$ は、時刻 $(t-1)$ におけるフレーム画像上の位置を表す2次元座標(フレーム画像内の、水平方向座標および垂直方向座標)のベクトルである。また、 $p(t)$ は、同様に、時刻 t におけるフレーム画像上の位置を表す2次元座標のベクトルである。マトリックス globalM は、時刻 $(t-1)$ から時刻 t までの間のグローバル動きを表す2行2列のマトリックスである。

【0052】

一方、ローカル動きの予測として、映像解析処理部20は、対象の顔領域内での顔の動きを予測する。ローカル動きの予測のために、映像解析処理部20は、カルマンフィルタの枠組みを用いる。ローカル動きの予測の処理手順は、例えば前記の非特許文献2にも開示されている。ローカル動きの予測のモデルは領域ごとに異なる。領域 r のローカル動きのモデルを $kf(r)$ と表現する。カルマンフィルタの予測関数 $kf(r).\text{pred}$ を用いた、領域 r の時刻 $(t-1)$ から時刻 t までの動き補償は、下の式(2)を用いて計算可能である。

【0053】

$$r(t) = kf(r).\text{pred}(r(t-1)) \cdots (2)$$

【0054】

本実施形態では、映像解析処理部20は、上記のグローバル動き予測とローカル動き予測とを組み合わせ、従来手法よりも精密な動き予測を行う。グローバル動き予測とローカル動き予測とを組み合わせた時刻 $(t-1)$ から時刻 (t) までの動き予測は、下の式(3)で表わされる。

【0055】

$$r(t) = \text{integrateM}(r(t-1)) = kf(r).\text{pred}(\text{globalM} \cdot r(t-1)) \cdots (3)$$

【0056】

式(3)において、 integrateM は、グローバル動き予測とローカル動き予測とを統合した動き予測モデルを表す。

【0057】

顔追跡装置1の全体的な処理負荷を適切にするため、比較的処理時間のかかる顔認識処理(顔認識処理部40の処理)をローレートで行う場合がある。言い換えれば、顔追跡装置1は、動き予測の処理を相対的に高い頻度で行い、顔認識処理を相対的に低い頻度で行うことができる。仮に動き予測の処理の頻度を顔認識処理の頻度に合わせてローレートで行った場合には、動き予測の精度が低くなる可能性がある。しかしながら、本実施形態の顔追跡装置1は、上記の通り、顔認識処理を相対的にローレートで行いながら、動き予測処理を相対的にハイレートで行うことができる。これにより、顔認識処理のための装置への負荷を軽減しながら、動き予測の精度を高く維持することができる。言い換えれば、本実施形態の顔追跡装置1は、顔認識処理のレートと動き予測処理のレートとを異ならせることにより、処理コストを抑制しながら、動き予測の精度を維持することができる。

【0058】

言い換えれば、顔追跡装置1は、映像解析処理部20による映像解析処理を、顔認識処理部40による顔認識処理よりも高頻度で行う。そして、映像解析処理部20は、複数回の動き予測の処理の結果である動き予測結果の情報(動き予測モデルが持つ動き量)を蓄積する。そして、追跡処理部50は、映像解析処理部20から渡される蓄積された動き予測結果の情報に基づいて処理を行う。

【0059】

なお、追跡処理(追跡処理部50の処理)やデータ出力のための処理(出力整形処理部60の処理)のレートは、顔認識処理のレートに合わせられる。

【0060】

図2は、顔追跡装置1における、映像解析処理と追跡処理との時間的関係を示す概略図である。図示するように、顔追跡装置1が処理対象とする映像は、フレーム画像のシーケ

10

20

30

40

50

ンスである。映像は、例えば、30FPSあるいは60FPSなどといったレートでのフレーム画像のシーケンスである。なお、「FPS」は、「frames per second」（秒あたりフレーム枚数）の略である。映像解析処理部20は、これらのフレーム画像の各々について、映像解析処理を行い、映像解析結果を出力する。一方、追跡処理部50は、適宜定められる所定数の連続するフレーム画像を対象として、顔の追跡処理を行う。なお、追跡処理部50が追跡処理を行う際には、映像解析処理部20が出力する映像解析結果を利用する。

【0061】

図2における映像解析処理は、前述の動き予測処理を含むものである。また、追跡処理は、顔認識処理と同じ頻度で、顔認識処理の結果を利用して実行されるものである。図2に示す時間的關係で処理を行う場合、1回の顔認識処理および追跡処理に、複数回の動き予測処理が対応する。このとき、顔追跡装置1は、映像解析処理部20による動き予測処理で求められる動きパラメーターを蓄積しておくようにする。

【0062】

つまり、映像解析処理部20は、時刻 $(t-1)$ におけるフレーム画像と時刻 t におけるフレーム画像とに基づいて、時刻 t における映像解析結果を求め、出力する。以下、同様に、映像解析処理部20は、時間的に隣接する2つのフレーム画像に基づいて、映像解析結果を求める。つまり、映像解析処理部20は、第1の映像解析処理の結果として動き予測パラメーター M_1 （グローバル動きの予測パラメーター）を出力する。以下、同様に、映像解析処理部20は、第 k （ k は正整数）の映像解析処理の結果として動き予測パラメーター M_k （グローバル動きの予測パラメーター）を出力する。映像解析処理部20は、この各回の動き予測パラメーター M_k （ $k=1, 2, \dots$ ）を蓄積して、追跡処理用のグローバル動き予測パラメーター $GlobalM$ を計算する。また、映像解析処理部20は、それぞれの顔領域 r についてのローカル動きの予測モデルも、同様に蓄積し、追跡処理用のローカル動き予測パラメーターを計算する。映像解析処理部20は、これらの蓄積された動き予測パラメーターを、追跡処理のタイミングで、追跡処理部50に渡す。

【0063】

また、映像解析処理部20は、シーンチェンジ検出の処理も、同様に、ハイレートの頻度で行うことができる。この場合、第1の映像解析処理の結果として、 $is_sc_change_1$ （値は、前述の通り、0（シーンチェンジ無し）または1（シーンチェンジ有り））を求める。以下同様に、映像解析処理部20は、第 k （ k は正整数）の映像解析処理の結果として $is_sc_change_k$ を求める。そして、映像解析処理部20は、各回の映像解析処理の結果としてのシーンチェンジの有無の判定結果 $is_sc_change_k$ （ $k=1, 2, \dots$ ）に基づいて、追跡処理用の is_sc_change を求める。なお、 $k=1, 2, \dots$ の中のいずれかに1つにおいて $is_sc_change_k=1$ ならば、映像解析処理部20は、 $is_sc_change=1$ とする。また、 $k=1, 2, \dots$ の中のすべてにおいて $is_sc_change=0$ ならば、映像解析処理部20は、 $is_sc_change=0$ とする。映像解析処理部20は、蓄積されたシーンチェンジ判定結果である is_sc_change の値を、追跡処理のタイミングで、追跡処理部50に渡す。なお、 $is_sc_change=1$ の場合には、追跡処理部50は、動き予測パラメーターをリセットする。一旦シーンチェンジが検出されて動き予測パラメーターがリセットされた場合には、映像解析処理部20は、次の顔認識処理（および追跡処理）のタイミングまで、すべての動き予測処理を中止するようにしてもよい。

【0064】

映像解析処理部20が行うボケ判定は、処理対象の映像内の顔領域ごとに、その領域がボケの領域であるか否かを判定するものである。なお、映像解析処理部20は、顔領域がボケている場合にも、顔領域の検出を行う。ボケ領域の映像に基づいて顔認識処理を行うと、認識精度が悪くなる。つまり、ボケ顔画像を対象として顔認識処理を行っても、認識精度が低く、正しくない人物を認識してしまう可能性が高い。これは、顔追跡装置1全体の顔認識精度の低下を招いてしまう。そのような精度劣化要因を避けるため、本実施形態の顔追跡装置1は、領域ごとのボケ判定（ボケている領域であるか否かの判定）を行い、

ボケ領域である場合には顔認識処理を行わない。ただし、顔追跡装置 1 は、顔認識処理を行わない場合にも、顔領域の検出結果の情報については、追跡処理部 50 への引き渡しを行う。

【0065】

映像解析処理部 20 は、顔領域 r がボケ領域であるか否かを判定し、判定結果を変数 $is_blur(r)$ の値として設定する。 $is_blur(r) = 1$ は、領域 r がボケ画像であると判定されたことを表す。 $is_blur(r) = 0$ は、領域 r がボケ画像ではないと判定されたことを表す。映像解析処理部 20 は、それぞれの顔領域についてボケ領域であるか否かの判定を行う。

【0066】

映像解析処理部 20 は、一例として、顔領域 r の画像が含むエッジ成分を解析することによってボケの有無を判定する。この方法自体は、従来技術を用いて実現可能である。具体的には、映像解析処理部 20 は、顔領域 r 内のエッジ成分を検出し、そのエッジ成分のシャープさの度合いが所定レベル以下である（その部分の画像が高い周波数成分を持たない）場合に「ボケ有り」（ $is_blur(r) = 1$ ）と判定する。また、映像解析処理部 20 は、顔領域 r 内のエッジ成分のシャープさの度合いが所定レベル以上である（その部分の画像が高い周波数成分を持つ）場合に「ボケ無し」（ $is_blur(r) = 0$ ）と判定する。

【0067】

映像解析処理部 20 は、別の方法でボケ画像の判定を行うようにしてもよい。この手法では、映像解析処理部 20 は、顔認識処理部 40 が顔認識処理を行う際に同時にボケ画像の有無の判定を行う。即ち、この手法を用いる場合、顔データベース 30 は、多数のボケ顔画像から抽出した画像の特徴量を、ボケ画像の人物 ID に関連付けて、予め保持しておく。これにより、顔の認識処理の対象の顔領域 r の特徴量を、顔データベース 30 が持つ特徴量と照合したときに、当該顔領域 r がボケ領域であることを表す。つまり、映像解析処理部 20 によるボケ画像の有無の判定と、顔認識処理部 40 による顔認識の処理とを一体の処理として行うことができる。つまり、顔追跡装置 1 全体の計算量を削減することが可能となる。

【0068】

図 3 は、顔認識処理と一体化したボケ判定処理のための、顔データベースの構築方法を模式的に示す概略図である。図示するように、本手法では、多数のボケ顔の画像を準備する。そして、顔追跡装置 1 は、これらのボケ顔の画像の特徴量を抽出する。画像の特徴量を抽出する処理は、顔認識処理部 40 が持つ機能を用いて行うことができる。そして、顔追跡装置 1 は、得られたボケ顔特徴量のデータを、ボケ顔を表す人物 ID に関連付けて、顔データベース 30 に登録する。なお、不特定多数の人のボケ顔に対して、ボケ顔を表す人物 ID を 1 つ与えれば十分である。

【0069】

つまり、映像解析処理部 20 は、顔領域がボケ領域であるか否かの判定を行う。追跡処理部 50 は、その判定結果に基づき、ボケ領域である顔領域に関しては、顔領域のトラックを更新する際には認識スコアの値を最低値（最悪スコア値）とする。ボケ判定の一つの手法は、次の通りである。顔認識処理部 40 は、顔領域がボケ領域である場合の特徴量と、当該ボケ領域の特徴量に関連付けられた特殊な人物 ID 識別情報と、を関連付けた情報（顔データベース 30 等）にアクセス可能と予めしておく。映像解析処理部 20 は、顔認識処理部 40 が顔領域についての顔認識処理を行った結果としてボケ領域の特徴量に関連付けられた特殊な人物 ID を同定した場合に、当該顔領域がボケ領域であると判定する、ようにできる。

【0070】

追跡処理部 50 による追跡処理の詳細は、下記の通りである。追跡処理部 50 は、顔認識処理部 40 が出力する結果と、映像解析処理部 20 が出力する結果とを用いて、時間軸上での対応関係から、同一人物の顔領域のトラックを求める処理を行う。追跡処理部 50 のこの処理によって、同一人物の顔領域のトラックが映像中で同定できれば、たとえある時刻のフレーム画像において顔認識処理結果の信頼性が低くても、同一トラックに属する

10

20

30

40

50

他の時刻のフレーム画像で信頼性の高い顔認識処理結果を得られれば、置き換えにより、結果として精度の高い顔認識を実現することができる。

【 0 0 7 1 】

図 4 は、追跡処理部 5 0 による追跡処理の手順を示すフローチャートである。以下、このフローチャートに沿って説明する。なお、このフローチャートは、ある 1 時点での追跡処理の手順を示すものである。つまり、追跡処理部 5 0 は、追跡処理を行うタイミングごとに、図 4 に示す処理を実行する。

【 0 0 7 2 】

時刻 t における追跡処理部 5 0 の処理の前提となるデータは次の通りである。追跡処理部 5 0 は、顔認識処理部 4 0 と映像解析処理部 2 0 からの出力信号（データ）を取得する。顔認識処理部 4 0 から追跡処理部 5 0 へは、検出した顔領域（フレーム画像内の位置情報）や、顔認識処理の結果である認識スコアや、認識結果である人物 ID や、上記顔領域の特徴量が渡される。映像解析処理部 2 0 から追跡処理部 5 0 へは、シーンチェンジ検出信号 is_sc_change や、顔領域の統合動き予測パラメータ $integrateM$ が渡される。

【 0 0 7 3 】

また、追跡処理部 5 0 は、前回の追跡処理時、即ち時刻 $(t - 1)$ において追跡できているトラックの集合を保持している。時刻 $(t - 1)$ におけるトラックの集合は、下の式 (4) の通りである。なお、式 (4) において n_Tr は、当該トラックの集合に含まれるトラックの数である。

【 0 0 7 4 】

$Tr(t-1) = [tr[0](t-1), tr[1](t-1), \dots, tr[n_Tr-1](t-1)] \dots (4)$

【 0 0 7 5 】

なお、各トラック $tr[k](t-1)$ ($0 \leq k \leq n_Tr - 1$) は、データの構造体であり、その構造体の要素は下記の $bbox$ 、 $feat$ 、 $recog_ids$ 、 $recog_scores$ を含む。

$bbox$ は、時刻 $(t - 1)$ におけるトラックに対応する顔領域の矩形領域情報（フレーム画像内の位置情報）である。

$feat$ は、時刻 $(t - 1)$ のトラックにおける顔領域の特徴量である。

$recog_ids$ は、時刻 $(t - 1)$ の当該トラックに対応する顔認識結果の人物 ID の候補である。人物 ID の候補が複数であってもよい。複数の人物 ID の候補は、例えば、スコアの高い順に並べられている。

$recog_scores$ は、時刻 $(t - 1)$ の当該トラックにおける上記人物 ID 候補に対応するスコアの値である。上記の人物 ID が並べられている順と同じ順（例えばスコアの高い順）にソートされている。

【 0 0 7 6 】

構造体の要素にアクセスする際の記法は、次の通りである。例えば、時刻 $(t - 1)$ におけるトラック集合全体についての $bbox$ （矩形領域情報）は、 $Tr(t-1).bbox$ と表わされる。例えば、時刻 $(t - 1)$ における個別トラックについての $bbox$ は、 $tr[k](t-1).bbox$ と表わされる。但し、 k は、トラックの集合内における個別トラックの指標値であり、 $0 \leq k \leq n_Tr - 1$ である。

【 0 0 7 7 】

なお、追跡処理部 5 0 の初回の処理の前（即ち、時刻 0 のとき）には、トラック集合の初期化処理が行われる。つまり、 $Tr(0) = []$ （空集合）とする初期化が行われる。なお、シーンチェンジによってトラックをリセットする場合にも、トラック集合は空集合に再設定される。

【 0 0 7 8 】

また、追跡処理部 5 0 が顔認識処理部 4 0 から受け取る時刻 t の顔検出領域および顔認識結果の情報は、下の式 (5) の通りである。なお、式 (5) において、 n_Det は、時刻 t の顔認識処理における顔検出領域の数（即ち、顔認識の数）である。

【 0 0 7 9 】

$Det(t) = [det[0](t), det[1](t), \dots, det[n_Det-1](t)] \dots (5)$

【 0 0 8 0 】

トラックの場合と同様に、各顔検出領域 $\text{det}[k](t)$ ($0 \leq k \leq n_{\text{Det}} - 1$) は、データの構造体であり、その構造体の要素は下記の bbox 、 feat 、 recog_ids 、 recog_scores を含む。

bbox は、時刻 t における顔検出領域の矩形領域情報（フレーム画像内の位置情報）である。

feat は、時刻 t における上記顔検出領域の特徴量である。

recog_ids は、時刻 t の当該顔検出領域に対応する顔認識結果の人物 ID の候補である。人物 ID の候補が複数であってもよい。複数の人物 ID の候補は、例えば、スコアの高い順に並べられている。

recog_scores は、時刻 t の当該顔検出領域における上記人物 ID 候補に対応するスコアの値である。上記の人物 ID が並べられている順と同じ順（例えばスコアの高い順）にソートされている。

【 0 0 8 1 】

構造体の要素にアクセスする際の記法は、次の通りである。例えば、時刻 t おける顔検出領域の集合全体についての bbox （矩形領域情報）は、 $\text{Det}(t).\text{bbox}$ と表わされる。例えば、時刻 t における個別の顔検出領域についての bbox は、 $\text{det}[k](t).\text{bbox}$ と表わされる。但し、 k は、顔検出領域の集合内における個別の顔検出領域の指標値であり、 $0 \leq k \leq n_{\text{Det}} - 1$ である。

【 0 0 8 2 】

追跡処理部 50 が扱うデータは、さらに、矩形領域間の重なり度合いを表すマトリックスと、特徴量間の距離を表すマトリックスとを含む。

【 0 0 8 3 】

矩形領域間の重なり度合いを表すマトリックスは、 $\text{IoU}(\text{bboxA}, \text{bboxB})$ と表わされる。 IoU は、「Intersection over Union」（領域同士の和集合に対する、当該領域同士の積集合の比率）を表す。追跡処理部 50 は、トラックが持つ矩形領域（上記の bboxA ）と、顔検出領域が持つ矩形領域（上記の bboxB ）との重なり度合いを求め、追跡処理に利用する。このマトリックスは、 $n_{\text{bboxA}} \times n_{\text{bboxB}}$ のサイズを持つ 2 次元のマトリックスである。ただし、 n_{bboxA} は bboxA の領域の数であり、 n_{bboxB} は bboxB の領域の数である。このマトリックスの要素の値が 1 のとき、対応する領域が完全に一致する。要素の値が正の値（0 以上且つ 1 以下）の場合には、値が小さいほど、対応する領域間の重なり度合いが小さくなっていく。要素の値が負の値である場合には、対応する領域は重なる部分を全く持たず、その値の絶対値が大きいほど、両領域間の距離は離れている。

【 0 0 8 4 】

特徴量間の距離を表すマトリックスは、 $\text{Dis}(\text{FeatA}, \text{FeatB})$ と表わされる。追跡処理部 50 は、トラックが持つ矩形領域の特徴量（上記の FeatA ）と、顔検出領域が持つ矩形領域の特徴量（上記の FeatB ）との距離を求め、追跡処理に利用する。このマトリックスは、 $n_{\text{FeatA}} \times n_{\text{FeatB}}$ のサイズを持つ 2 次元のマトリックスである。ただし、 n_{FeatA} は FeatA の特徴量の数（トラックが持つ矩形領域の数と同一）であり、 n_{FeatB} は FeatB の特徴量の数（顔検出領域が持つ矩形領域の数と同一）である。距離の値は、0 または正の値である。2 つの特徴量が同一である場合には、当該 2 つの特徴量間の距離は 0 である。2 つの特徴量が似ていない度合いが高いほど、当該特徴量間の距離は大きくなる。

【 0 0 8 5 】

追跡処理部 50 は、以上のデータに基づいて追跡処理を行う。言い換えれば、追跡処理部 50 が行う追跡処理は、時刻 ($t - 1$) におけるトラック集合 $\text{Tr}(t-1)$ と、時刻 t における顔検出領域の集合 $\text{Det}(t)$ との対応付けの問題に還元される。追跡処理部 50 は、時刻の経過（例えば、時刻 ($t - 1$) から時刻 t に）に伴って図 4 に示す処理を 1 回ずつ行い、これを繰り返すことにより、トラック集合 $\text{Tr}(t)$ のデータを順次更新していく。

【 0 0 8 6 】

追跡処理部 50 の時刻 t における処理手順は次の通りである。まず、ステップ S 1 1 に

10

20

30

40

50

において、追跡処理部50は、映像解析処理部20から渡される情報に基づいて、前回の追跡処理のときからシーンチェンジがあったか否かを判定する。シーンチェンジの有無は、変数is_sc_changeの値によって示される。シーンチェンジがあったとき、即ちis_sc_changeの値が1のとき(ステップS11: YES)には、ステップS21に飛ぶ。シーンチェンジがなかったとき、即ちis_sc_changeの値が0のとき(ステップS11: NO)には、時刻(t-1)におけるトラック集合を現トラック集合として(existTr=Tr(t-1))、次のステップS12に進む。

【0087】

次に、ステップS12において、追跡処理部50は、動き補正処理を行う。つまり、追跡処理部50は、現トラック集合existTr.bboxが持つ各トラックに対して、映像解析処理部20から渡された統合動き予測モデルintegrateMによる動き補正の処理を行う。

10

【0088】

次に、ステップS13において、追跡処理部50は、領域の重なり度を計算する。具体的には、追跡処理部50は、ステップS12における動き補正処理の結果として得られる矩形領域を、sc倍に拡張する。この拡張は、動き補正処理が含み得る動き補正の誤差をカバーするためのものである。拡張倍率scの値は、1.0以上の、適宜定める値とする。例えば、1.5、sc=2.5などとしてよい。本実施形態では、一例としてsc=2.0とする。この拡張により、矩形領域の中心位置は変わらず、矩形領域の大きさがsc倍となる。そして、追跡処理部は、拡張後の矩形領域と、顔認識処理部40から渡されたDet(t).bboxの重なり度合いを計算する。そして、追跡処理部50は、重なり度合いのマ

20

【0089】

つまり、追跡処理部50は、前記第1位置情報と前記第2位置情報との重なり度合いに基づいて、前記トラックを更新する処理を行う。

【0090】

次に、ステップS14において、追跡処理部50は、現トラックが持つ顔領域の特徴量と、顔認識処理部40から渡された時刻tにおける顔領域の特徴量とを比較する。つまり、追跡処理部50は、現トラックにおける特徴量であるExistTr.featsと、顔認識処理部40から渡された特徴量であるDet(t).featsとの、特徴量間の距離を表すマトリックスDisMtxを求める。

30

【0091】

次に、ステップS15において、追跡処理部50は、領域フィルター処理を行う。つまり、追跡処理部は、ステップS13で求めたGateMtxを用いて、ステップS14で求めたマトリックスDisMtxに含まれる無効成分をフィルターアウトする。言い換えれば、ステップS13における閾値処理の結果として、トラックが持つ顔領域と、顔検出領域が持つ顔領域とが、相互に全く重ならない領域として扱われる場合には、追跡処理部50がそのトラックとその顔検出領域とを対応付けることはない。

40

【0092】

次に、ステップS16において、追跡処理部50は、ステップS15の結果として得られたDisMtxに基づいて、時刻(t-1)までの顔領域トラックと時刻tでの顔検出領域との対応関係を判定する。言い換えれば、追跡処理部50は、領域間の重なり度合い(ただし、閾値処理済み)と、各領域の特徴量間の距離とに基づいて、時刻(t-1)のけるトラックと時刻tにおける顔検出領域との対応付けを行う。なお、本ステップにおいて、追跡処理部50は、例えばハンガリアンアルゴリズム(Hungarian algorithm、ハンガリー法)を用いて上記の対応関係を求める。ハンガリアンアルゴリズム自体は、既存技術によるものである。言い換えれば、追跡処理部50は、顔領域トラックの番号(tr_no)と、顔検出領域の番号(det_no)との対応付けを試みる。追跡処理部50は、求められた対

50

応関係に応じて、顔領域の処理を次の通り分岐させる。顔領域トラックと顔検出領域とが対応している場合 (matched_pairs) には、ステップ S 1 7 に進む。顔検出領域に対応する顔領域トラックが存在しない場合 (unmatched_dets) には、ステップ S 1 8 に進む。顔領域トラックに対応する顔検出領域が存在しない場合 (unmatched_tracks) には、ステップ S 1 9 に進む。なお、上記 matched_pairs は、顔領域トラックの番号と顔検出領域の番号との組合せの集合である (matched_pairs {tr_no,det_no})。また、上記 unmatched_dets は、顔検出領域の番号の集合である (unmatched_dets {det_no})。また、unmatched_tracks は、顔領域トラックの番号の集合である (unmatched_tracks {tr_no})。

【 0 0 9 3 】

10

なお、追跡処理部 5 0 は、ステップ S 1 6 における条件分岐を、領域ごと (matched_pair ごと、unmatched_det ごと、そして unmatched_tracks ごと) に行う。そして、追跡処理部 5 0 は、ステップ S 1 6 で分類された領域ごとに、次のステップ S 1 7、S 1 8、S 1 9 のそれぞれの処理を行う。

【 0 0 9 4 】

次に、ステップ S 1 7 において、追跡処理部 5 0 は、該当する顔検出領域 (Detection) の情報を、対応する既存の顔領域トラック (Track) に追加する。

【 0 0 9 5 】

次に、ステップ S 1 8 において、追跡処理部 5 0 は、該当する顔検出領域 (unmatched detection) に基づいて新しい顔領域トラック (Track) を登録する。つまり、追跡処理部 5 0 は、その顔検出領域を、新たなトラックとして、既存トラック集合に追加する。

20

【 0 0 9 6 】

次に、ステップ S 1 9 において、追跡処理部 5 0 は、該当する顔領域トラック (unmatched track) を、喪失された (lost) トラック、あるいは削除された (deleted) トラックとして登録する。具体的には、追跡処理部 5 0 は、既存トラックに対応する顔検出領域がなくなった時点でその既存トラックを喪失状態 (lost) として、所定期間この喪失状態が続く場合には、そのトラックを終了 (削除、deleted) として、以後の追跡処理の対象から外してよい。

【 0 0 9 7 】

すべての領域について上記ステップ S 1 7、S 1 8、S 1 9 の処理が完了すると、次にステップ S 2 0 において、追跡処理部 5 0 は、Tracks を更新する。つまり、追跡処理部 5 0 は、ステップ S 1 6 で判定した対応関係に基づいて、トラックを更新し、時刻 t におけるトラック集合 $Tr(t)$ を生成する。

30

【 0 0 9 8 】

なお、トラックを更新する際に、領域 $tr[i](t)$ がボケ顔領域である ($is_blur = 1$) である場合には、追跡処理部 5 0 は、当該領域の認識スコアを強制的に最低値 (通常は、0.0) とする。即ち、追跡処理部 5 0 は、そのような領域について、 $tr[i](t).recog_score = 0.0$ とする。

【 0 0 9 9 】

一方、ステップ S 2 1 に進んだ場合 (ステップ S 1 1 の判定においてシーンチェンジ有だった場合) には、追跡処理部は、トラックをリセットする。即ち、時刻 t におけるトラックを $Tr(t) = []$ (空集合) とする。つまり、この場合、 $Tr(t-1)$ までのトラック情報は以後において使用されなくなる。

40

【 0 1 0 0 】

つまり、追跡処理部 5 0 は、シーンチェンジが検出された場合には、トラックを初期状態にリセットする。

【 0 1 0 1 】

以上説明した図 4 のフローチャートの処理を繰り返した結果として、追跡処理部 5 0 は、顔認識結果の情報を持つ複数の顔領域トラックを、出力する。追跡処理部 5 0 は、この複数の顔領域トラックのデータを、出力整形処理部 6 0 に渡す。

50

【 0 1 0 2 】

出力整形処理部 6 0 は、追跡処理部 5 0 から渡される追跡処理の結果のデータを整形し、出力する。具体的には、出力整形処理部 6 0 は、追跡処理部 5 0 が出力したトラック（図 4 の追跡処理による更新後のトラック）が持つ各フレーム画像での顔認識処理の結果を、認識スコアの高い順に出力するように整形する。時刻 t におけるトラック $tr[k]$ （ k は、特定のトラックに対する指標値）を例にとると、次の通りである。

【 0 1 0 3 】

時刻（ $t - 1$ ）までにおいては、当該トラックは、顔認識に関する下記のデータを含む。

$tr[k](t-1).recog_scores$ は、時刻（ $t - 1$ ）までの認識スコアを並べたものである。ここでは、認識スコアは、高い順から（即ち降順に）並んでいるものとする。

$tr[k](t-1).recog_ids$ は、上記の $tr[k](t-1).recog_scores$ に対応する人物 ID の並びである。つまり、 $tr[k](t-1).recog_ids$ は、当該トラックの時刻（ $t - 1$ ）までの、認識スコアの降順に並べられた人物 ID（認識結果）である。

【 0 1 0 4 】

ここで、時刻 t における顔検出領域 $det[j]$ が、トラック $tr[k]$ に追加される場合を想定する。すると、顔検出領域 $det[j]$ に関する顔認識結果である $det[j].recog_scores$ および $det[j].recog_ids$ も、認識結果に関する情報としてトラック $tr[k]$ に追加されることになる。このとき、トラック $tr[k]$ が元々持っていた認識スコアおよび人物 ID（認識結果）の情報に加えて、新たに追加される情報も含めて、認識スコアの降順でのソーティングが行われる。このソーティングを、出力整形処理部 6 0 が行う。この処理は、下の式（6）および式（7）で表わされる。

【 0 1 0 5 】

```
tr[k](t).recog_scores =  
sort([tr[k](t-1).recog_scores+det[j].recog_scores],topK) ・・・(6)  
tr[k](t).recog_ids =  
argsort([tr[k](t-1).recog_ids+det[j].recog_ids],topK) ・・・(7)
```

【 0 1 0 6 】

なお、式（6）に示す処理 $sort(listA, topK)$ は、値のリスト $listA$ の要素を、その値順に（ここでは降順に）、 $topK$ 文だけ並べ替える処理を表す。なお、式（6）内の演算子「+」は、リスト同士の連結を行う演算子である。つまり、式（6）は、認識スコアのリスト $tr[k](t-1).recog_scores$ と認識スコアのリスト $det[j].recog_scores$ とを連結して成るリストに対するソート処理を表す。

【 0 1 0 7 】

また、式（7）に示す処理 $argsort(listA, topK)$ は、対応する $recog_scores$ （認識スコア）をソートキーとして、 $listA$ を降順に、 $topK$ 文だけ並べ替える処理を表す。ここでも、演算子「+」はリスト同士の連結を行う演算子である。つまり、式（7）は、式（6）に対応する順に、人物 ID のリスト $tr[k](t-1).recog_ids$ および人物 ID のリスト $det[j].recog_ids$ をソートする処理を表す。

【 0 1 0 8 】

このように、本実施形態では、あるトラックに関して、時間軸にわたって、ある人物 ID についての認識スコアの最大値が残るように、トラックの更新の処理を繰り返す。これにより、ある顔領域のスコアが低い時間帯（顔領域がボケ領域である時間帯を含む）の認識結果に左右されずに、顔領域のスコアが高い時間帯の認識結果に基づいて、トラックの顔認識を行うことができるようになる。

【 0 1 0 9 】

つまり、顔追跡装置 1 は、トラック単位で、認識スコアの高い順にソートされた顔認識結果のデータ（人物 ID および認識スコア）を保持し、更新する。これにより、処理対象の映像中に登場する人物に対して、安定した認識結果を得ることが可能となる。

【 0 1 1 0 】

図 5 は、あるトラックについての認識スコアの時間推移の例を示す概略図である。図示するように、このトラックは、認識結果の候補として、人物 A および人物 B の人物 ID を持つものである。当該トラックに関して、人物 A についてのスコアは、時刻 t_1 において 0.5 であり、時刻 t_2 において 0.8 である。一方、人物 B についてのスコアは、時刻 t_1 において 0.6 であり、時刻 t_2 において 0.6 である。

【0111】

このようなスコアが生じるのは、例えば、このトラックについての顔認識の正解が「人物 A」であり、時刻 t_1 においては人物 A が横を向いていたことによって「人物 A」の認識スコアが 0.5 に低下し、時刻 t_2 においては人物 A が正面を向いていたことによって認識スコアが 0.8 に上昇している場合である。つまりこの場合、時刻 t_1 においては、正解であるべき人物 A のスコアが、不正解である人物 B のスコアよりも低くなってしまっている。なお、ここでは $t_1 < t_2$ である。

【0112】

上記の場合、時刻 t_1 において、当該トラックが持つデータは、次の通りである。

```
tr(t1).recog_socres = [0.6,0.5]
```

```
tr(t1).recog_ids = [B,A]
```

また、時刻 t_2 において、当該トラックが持つデータは、次の通りである。

```
tr(t2).recog_scores = [0.8,0.5]
```

```
tr(t2).recog_ids = [A,B]
```

【0113】

時刻 t_1 における認識スコアと、時刻 t_2 における認識スコアとを統合すると、次の通りである。

```
tr.recog_score = [0.8,0.6]
```

```
tr.recog_ids = [A,B]
```

つまり、時間経過を通したトラックの認識結果としては、顔追跡装置 1 は、人物 A の、よりスコアの高い正面顔のときの状態を優先して用いていることとなる。これにより、顔追跡装置は、正解である人物 A を、人物 B よりも上位の認識結果として出力する。このように、顔追跡装置 1 は、時間経過を通したトラックの期間中において最も認識スコアの高い人物を、認識結果とする。これにより、顔追跡装置 1 は、より信頼性の高い認識処理を実現することが可能となる。

【0114】

なお、前述の通り、ボケ検出結果 `is_blur` が 1 のときには、顔追跡装置 1 は、当該トラックのその時刻における顔認識スコアを強制的 0.0 とする。つまり、トラックの顔領域がボケているときの認識結果は、時間を通したトラック全体の認識スコアの計算に影響を与えない。

【0115】

以上説明したように、顔追跡装置 1 は、映像内に出現する顔を追跡しながら、その顔の人物 ID を特定する。これにより、顔追跡装置 1 は、トラックごとの、人物 ID の候補と、その人物 ID に対応付けられたスコアとのデータを求め、出力することができる。

【0116】

図 6 は、顔追跡装置 1 による追跡処理の結果として得られるデータ構成の一例を示す概略図である。図示するように、顔追跡装置 1 は、特定の映像コンテンツに関して、トラックを実体とするデータの集合を出力することができる。図示するトラックの主キーは、例えば、トラック ID である。処理対象とした映像コンテンツは、ユニークなコンテンツ ID を持つことができる。つまり、1 つのコンテンツ ID は、複数の（多数の）トラックに関連付けられる。トラックの属性は、開始時刻（開始フレーム）、終了時刻（終了フレーム）、認識スコア（リスト）、人物 ID（リスト）を含む。開始時刻および終了時刻のそれぞれは、当該映像コンテンツにおける相対時刻である。開始時刻は、開始フレームに対応する。終了時刻は、終了フレームに対応する。各トラックの、開始時刻および終了時刻の精度は、追跡処理を行う頻度に依存する（図 2 参照）。つまり、追跡処理を十分に頻繁

10

20

30

40

50

に（例えば、１秒に１回、あるいは０．５秒に１回等）行うことにより、トラックの開始時刻および終了時刻の精度は、実用上十分なものとなる。ただし、追跡処理の頻度についての制約はない。認識スコアは、人物ＩＤに対応するスコアの値のリストである。認識スコアは、スコアが降順に並べられたリストのデータである。人物ＩＤは、映像内に登場する人物のリストである。人物ＩＤのソート順は、対応する認識スコアの降順である。このように、人物ＩＤをキーとしてトラックの集合のデータを検索することにより、該当する人物が登場する映像内のシーン（開始時刻および終了時刻）を、所定の精度で得ることが可能となる。

【０１１７】

図６内に示すように、人物ＩＤは、人物をユニークに特定するための識別情報である。人物ＩＤは、１つまたは複数の人物名（表記）のデータに関連付けられる。また、人物ＩＤは、１つまたは複数のサムネイル画像のデータに関連付けられる。つまり、人物名表記から、人物ＩＤを求めることが可能である。また、サムネイル画像が選択されると、人物ＩＤを求めることが可能である。

【０１１８】

図７は、人物名あるいはサムネイル画像を起点として、当該人物が登場する映像コンテンツのシーンを取り出すための情報経路の例を示す概略図である。図６を参照しながら説明した通り、例えば人物名表記が文字列として入力されたり、提示された多数のサムネイル画像（人物の顔の画像）から特定のサムネイル画像が選択されたりすると、人物ＩＤを特定することが可能である。人物ＩＤが得られると、図６に示したトラックの集合のデータを検索することにより、当該人物ＩＤに対応する、映像コンテンツのコンテンツＩＤと、その相対位置（開始時刻（開始フレーム）および終了時刻（終了フレーム））の情報を、１セットまたは複数セット、特定することができる。これにより、当該人物ＩＤに対応する映像シーンを、利用者等に対して提示することが可能となる。

【０１１９】

図８は、図７で説明した映像シーンの検索のためのＧＵＩ（グラフィカルユーザーインターフェース）の一例を示す概略図である。言い換えれば、同図は、顔認識結果に基づく映像検索システムのＧＵＩの例を示す。なお、この映像検索システムは、本実施形態の顔追跡装置１による処理の結果として得られたトラックのデータを用いて、映像を検索することができる。図示する画面は、ＰＣや、タブレット端末や、スマートフォン等の携帯型端末の表示画面である。同図において、７０１は、映像表示領域である。また、７０２は、人物名表示領域である。また、７０３は、サムネイル画像表示領域である。また、７０４は、シーン特定情報表示領域である。本システムの利用者は、人物名表示領域７０２に表示されている人物名（文字列）の一つ、あるいは、サムネイル画像表示領域７０３に表示されているサムネイル画像（顔画像）の一つを選択してクリック（あるいはタップ）することができる。その結果、映像検索システムは、図７に示したように、当該人物に該当するトラックを抽出することができる。そして、映像検索システムは、抽出されたトラックの情報に基づき、特定の映像コンテンツの特定のシーン（相対位置）を、映像表示領域７０１に表示する。選択された人物に関連するトラックが複数存在する場合には、例えば、当該映像検索システムは、認識スコアの値に基づいて、当該人物に関して高いスコアを有するトラックを選択してもよい。

【０１２０】

なお、顔追跡装置１自身が検索処理部を備えて、図８に示す映像検索を行うようにしてもよい。その場合、検索処理部は、ＧＵＩの画面等から、人物ＩＤ（人物識別情報）と、人物ＩＤに関連付けられた人物名表記文字列と、人物ＩＤに関連付けられたサムネイル画像の指定と、のいずれかの検索キーを取得する。検索処理部は、取得した検索キーに基づいて、出力整形処理部６０が出力したトラックの集合のデータを検索することによって、特定の人物ＩＤに関連付けられた単数または複数のトラックを特定する。これによって特定されたトラックは、コンテンツＩＤと映像コンテンツ内の相対位置の情報とを持っている。よって、検索処理部は、トラックによって特定される映像シーンを、記憶装置等から

10

20

30

40

50

取得し、デコードし、映像として画面（映像表示領域 701）に表示することができる。

【0121】

図9は、顔追跡装置の内部構成の例を示すブロック図である。顔追跡装置1は、コンピュータを用いて実現され得る。図示するように、そのコンピュータは、中央処理装置901と、RAM902と、入出力ポート903と、入出力デバイス904や905等と、バス906と、を含んで構成される。コンピュータ自体は、既存技術を用いて実現可能である。中央処理装置901は、RAM902等から読み込んだプログラムに含まれる命令を実行する。中央処理装置901は、各命令にしたがって、RAM902にデータを書き込んだり、RAM902からデータを読み出したり、算術演算や論理演算を行ったりする。RAM902は、データやプログラムを記憶する。RAM902に含まれる各要素は、アドレスを持ち、アドレスを用いてアクセスされ得るものである。なお、RAMは、「ランダムアクセスメモリー」の略である。入出力ポート903は、中央処理装置901が外部の入出力デバイス等とデータのやり取りを行うためのポートである。入出力デバイス904や905は、入出力デバイスである。入出力デバイス904や905は、入出力ポート903を介して中央処理装置901との間でデータをやりとりする。バス906は、コンピュータ内部で使用される共通の通信路である。例えば、中央処理装置901は、バス906を介してRAM902のデータを読んだり書いたりする。また、例えば、中央処理装置901は、バス906を介して入出力ポートにアクセスする。

10

【0122】

なお、顔追跡装置1の少なくとも一部の機能をコンピュータで実現することができる。その場合、この機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現しても良い。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM、DVD-ROM、USBメモリー等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。つまり、「コンピュータ読み取り可能な記録媒体」とは、非一過性の（non-transitory）コンピュータ読み取り可能な記録媒体であってよい。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、一時的に、動的にプログラムを保持するもの、その場合のサーバーやクライアントとなるコンピュータシステム内部の揮発性メモリーのように、一定時間プログラムを保持しているものも含んでも良い。また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであっても良い。

20

30

【0123】

以上、実施形態を説明したが、本発明はさらに次のような変形例でも実施することが可能である。

【0124】

[第1変形例]

例えば、顔認識処理部40による顔認識処理は、ニューラルネットワーク等の機械学習手段を用いて行うようにしてもよい。その場合、顔データベース30が保持する人物IDごとの画像特徴量は、ニューラルネットワーク内の内部パラメーター値の集合として表わされ、保持されるものである。つまり、画像特徴量と人物IDとの関係の情報は、ニューラルネットワーク内の各ノードのパラメーター値として埋め込まれる。

40

【0125】

[第2変形例]

上記実施形態では、映像の過去側から未来側に向かって（順方向で）、顔領域を追跡し、トラックを更新する処理を行った。この方向が逆でもよい。即ち、顔追跡装置1は、映像の未来側から過去側に向かって（逆方向で）、顔領域を追跡し、トラックを更新するよ

50

うにしてもよい。また、顔追跡装置 1 が、順方向の追跡と、逆方向の追跡とを、組み合わせて処理するようにしてもよい。これにより、動画ファイル内の複数のトラックを開始から終了まで特定でき、そのトラックに対する認識結果を付与することが可能となる。

【 0 1 2 6 】

以上説明したように、本実施形態（その変形例を含む。以下において同様。）によれば、映像中に映っている人物の顔認識処理と、顔領域を追跡する処理とを組み合わせることによって、顔認識精度の向上および追跡精度の向上の両方を実現する。本実施形態の顔追跡装置 1 は、映像内の、時間方向にわたって人物が出現する領域に対して、一貫した認識結果を付与することが可能となる。

【 0 1 2 7 】

また、本実施形態は、顔の追跡処理の際に、シーンチェンジ（カット）の検出結果を利用する。つまり、本実施形態は、シーンチェンジが検出されると、追跡していたトラックを初期化する。これにより、例えば長時間の映像でシーンチェンジを含む場合にも、適切な追跡を行うことができる。言い換えれば、処理の混乱を生じないようにすることができる。

【 0 1 2 8 】

また、本実施形態は、認識処理の精度を下げる要因であるボケ画像に対応するために、積極的にボケ画像の検出処理を行い、ボケ画像が検出される場合にはその認識スコアを最低値にするなお、ボケによる認識への影響を軽減している。

【 0 1 2 9 】

また、本実施形態では、2種類の処理のレート（頻度）を変えて、それらの処理を組み合わせている。具体的には、映像解析処理部 20 による映像解析処理と、顔認識処理部 40 および追跡処理部 50 による認識および追跡処理とを、異なるレートで行えるようにしている。これにより、本実施形態は、2種類の動き予測処理の結果である動き予測を蓄積し、追跡処理時に利用している。これにより、装置全体での処理負荷を軽減しながら、必要な処理を高いレート（頻度）で実行することが可能となっている。

【 0 1 3 0 】

また、本実施形態は、顔領域の軌跡（トラック）の単位の詳細データを出力し、そのトラックに認識結果の詳細データを付与するようにしている。このような形式の詳細データを出力することにより、本実施形態は、トラック単位で、最も確からしい（単数または複数の）認識結果を与えることができるようになっている。このような出力データは、人物 ID をキーとして、その人物が出現している映像シーンを検索するのに向いている。これにより、人物を起点として映像検索を容易にすることができる。また、本実施形態は、そのような映像検索のための GUI の一例を提示した。

【 0 1 3 1 】

つまり、本実施形態は、映像内に出現する人物の顔認識を時間的な変化にも対応して一貫した形式で実施することを可能とする。認識結果を対象映像へのメタデータ付加に簡単に利用できるようなデータ構造を考案することにより、登場人物の検索を効果的に行える検索システムの開発が容易となる。

【 0 1 3 2 】

以上、この発明の実施形態について図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計等も含まれる。

【産業上の利用可能性】

【 0 1 3 3 】

本発明は、例えば、映像コンテンツ制作事業や、映像コンテンツ配信事業や、映像コンテンツ管理事業や、映像コンテンツ検索サービス事業等に利用することができる。但し、本発明の利用範囲はここに例示したものには限られない。

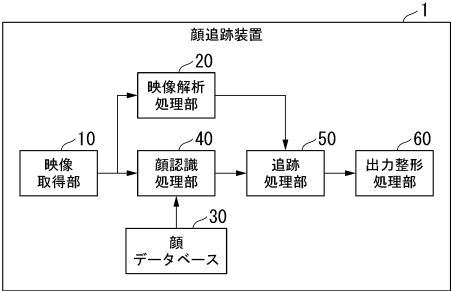
【符号の説明】

【 0 1 3 4 】

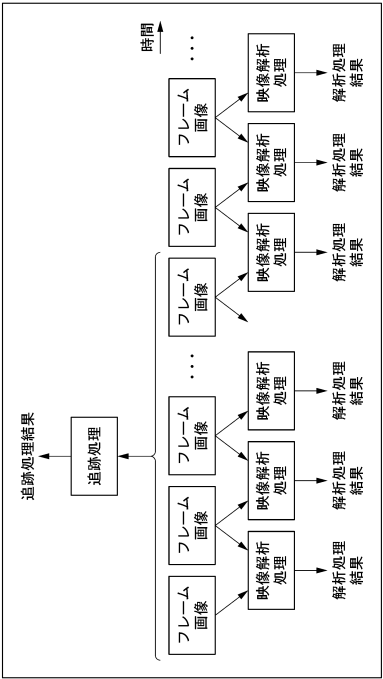
- 1 顔追跡装置
- 1 0 映像取得部
- 2 0 映像解析処理部
- 3 0 顔データベース
- 4 0 顔認識処理部
- 5 0 追跡処理部
- 6 0 出力整形処理部
- 7 0 1 映像表示領域
- 7 0 2 人物名表示領域
- 7 0 3 サムネール画像表示領域
- 7 0 4 シーン特定情報表示領域
- 9 0 1 中央処理装置
- 9 0 2 R A M
- 9 0 3 入出力ポート
- 9 0 4 , 9 0 5 入出力デバイス
- 9 0 6 バス

【図面】

【図 1】



【図 2】



10

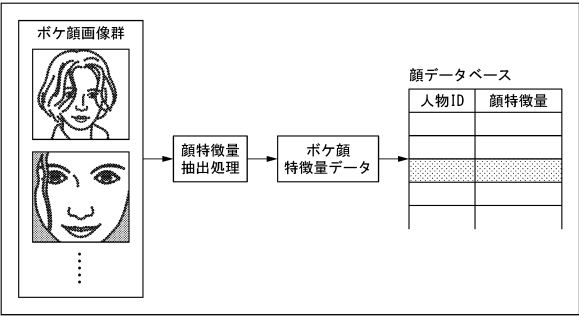
20

30

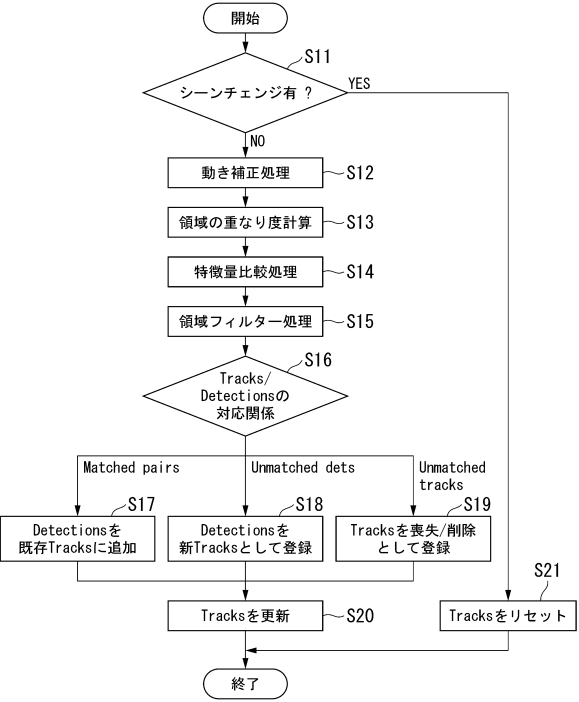
40

50

【図 3】



【図 4】

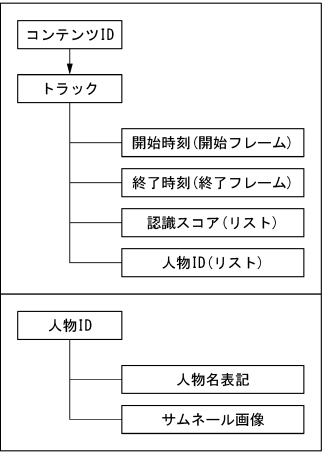


【図 5】

認識スコアの例

時刻	t1	t2	t1⇒t2のトラックとしてのスコア
人物A	0.5	0.8	0.8
人物B	0.6	0.6	0.6

【図 6】



10

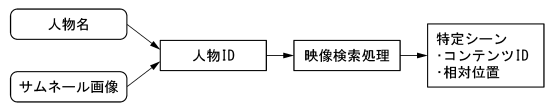
20

30

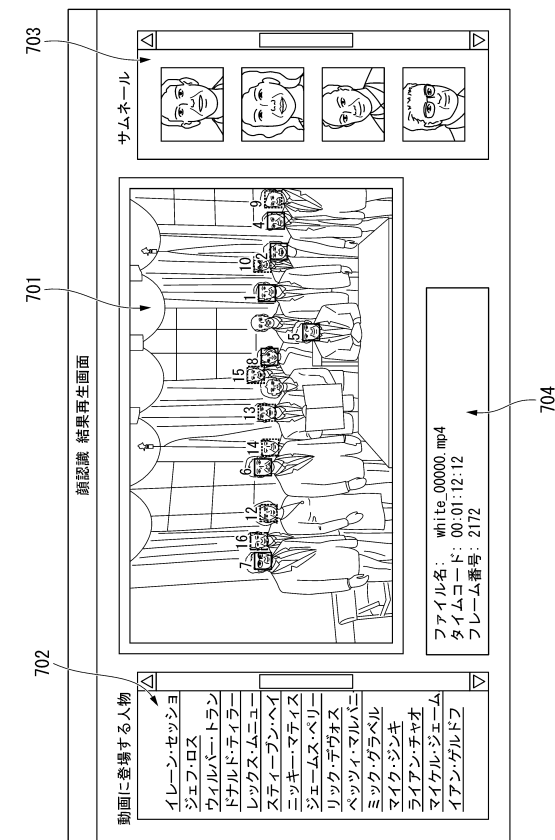
40

50

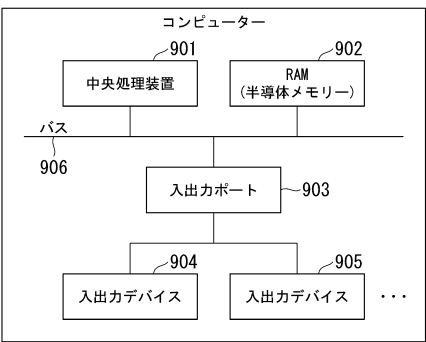
【図 7】



【図 8】



【図 9】



10

20

30

40

50

フロントページの続き

- 東京都世田谷区砧一丁目 1 0 番 1 1 号 日本放送協会放送技術研究所内
- (72)発明者 河合 吉彦
- 東京都世田谷区砧一丁目 1 0 番 1 1 号 日本放送協会放送技術研究所内
- (72)発明者 苗村 昌秀
- 東京都世田谷区砧一丁目 1 0 番 1 1 号 一般財団法人NHKエンジニアリングシステム内
- 審査官 村山 絢子
- (56)参考文献 特開 2 0 1 9 - 2 0 4 5 1 8 (J P , A)
- 特開 2 0 1 3 - 1 9 6 2 9 4 (J P , A)
- 特開 2 0 0 9 - 2 0 1 0 4 1 (J P , A)
- (58)調査した分野 (Int.Cl., D B 名)
- G 0 6 T 7 / 0 0 - 7 / 9 0
- H 0 4 N 7 / 1 8
- H 0 4 N 2 3 / 0 0 - 2 3 / 9 5 9
- H 0 4 N 5 / 2 2 2 - 5 / 2 5 7