

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第5899472号
(P5899472)

(45) 発行日 平成28年4月6日 (2016.4.6)

(24) 登録日 平成28年3月18日 (2016.3.18)

(51) Int.Cl.	F I
GO6T 7/00 (2006.01)	GO6T 7/00 350B
GO6T 1/00 (2006.01)	GO6T 1/00 340A
HO4N 7/18 (2006.01)	HO4N 7/18 R
GO8B 13/196 (2006.01)	GO8B 13/196

請求項の数 11 (全 23 頁)

(21) 出願番号	特願2012-117129 (P2012-117129)	(73) 特許権者	314012076 パナソニックIPマネジメント株式会社 大阪府大阪市中央区城見2丁目1番61号
(22) 出願日	平成24年5月23日 (2012.5.23)	(74) 代理人	100131451 弁理士 津田 理
(65) 公開番号	特開2013-242825 (P2013-242825A)	(74) 代理人	100106840 弁理士 森田 耕司
(43) 公開日	平成25年12月5日 (2013.12.5)	(74) 代理人	100113549 弁理士 鈴木 守
審査請求日	平成27年1月30日 (2015.1.30)	(72) 発明者	西村 純 福岡県福岡市博多区美野島4丁目1番62号 パナソニックシステムネットワークス株式会社内
		最終頁に続く	

(54) 【発明の名称】 人物属性推定システム、及び学習用データ生成装置

(57) 【特許請求の範囲】

【請求項1】

属性推定対象人物を撮影して画像を生成するカメラと、
前記カメラが生成した画像に写っている人物の属性を、推定モデルを用いて推定する属性推定部と、
前記カメラによる前記属性推定対象人物の撮影環境を示す、撮影環境データを取得する撮影環境データ取得部と、
人物画像である標準画像を取得する標準画像取得部と、
前記撮影環境データに応じて前記標準画像のデータを加工することにより、前記標準画像に前記撮影環境を反映させた疑似現場画像を生成する疑似現場画像生成部と、
前記疑似現場画像を用いて前記推定モデルの学習を行なう学習部と、
を備えたことを特徴とする人物属性推定システム。

【請求項2】

前記学習部は、前記カメラによる前記属性推定対象人物の撮影環境において、前記カメラで前記疑似現場画像が現出した被写体を撮影することにより得られた疑似現場画像撮影画像を学習用画像データとして用いて前記推定モデルの学習を行う
ことを特徴とする請求項1に記載の人物属性推定システム。

【請求項3】

前記学習部は、前記疑似現場画像生成部にて生成された前記疑似現場画像を学習用画像データとして用いて前記推定モデルの学習を行なう

ことを特徴とする請求項 1 に記載の人物属性推定システム。

【請求項 4】

前記標準画像には、被写体である人物の属性を示す属性データが対応付けられており、
前記学習部は、前記疑似現場画像の生成に用いられた前記標準画像に対応する属性データを、学習用正解属性データとして用いて前記推定モデルの学習を行う

ことを特徴とする請求項 1 から 3 のいずれかに記載の人物属性システム。

【請求項 5】

前記カメラで生成された画像に基づき、前記撮影環境データを算出する撮影環境推定部をさらに備え、

前記撮影環境データ取得部は、前記撮影環境推定部で算出された前記撮影環境データを取得する

ことを特徴とする請求項 1 から 4 のいずれかに記載の人物属性推定システム。

【請求項 6】

前記撮影環境データは、前記カメラにより前記属性推定対象人物が撮影される場所における照明状態を示すデータを含み、

前記疑似現場画像生成部は、前記照明状態を示すデータに応じて、前記標準画像を変換することにより、前記疑似現場画像を生成する

ことを特徴とする請求項 1 から 5 のいずれかに記載の人物属性推定システム。

【請求項 7】

前記属性推定部は、前記カメラが生成した画像に写っている人物の属性を、前記画像中の顔領域の部分画像に基づき推定するものであり、

前記撮影環境データは、前記属性推定対象人物が、前記カメラにより撮影される場合の顔の向きに関するデータを含み、

前記標準画像は、人物の顔を含む画像であって、

前記疑似現場画像生成部は、前記顔の向きに関するデータに応じて、前記標準画像の顔の向きを変換することにより、前記疑似現場画像を生成する

ことを特徴とする請求項 1 から 6 のいずれかに記載の人物属性推定システム。

【請求項 8】

前記撮影環境データは、前記カメラにより生成される画像内の 1 つ以上の代表的な人物検出領域ごとの撮影環境データである

ことを特徴とする請求項 1 から 7 のいずれかに記載の人物属性推定システム。

【請求項 9】

前記疑似現場画像生成部は、前記代表的な人物検出領域ごとの撮影環境データを用いて、前記代表的な人物検出領域ごとの疑似現場画像を生成し、

前記学習部は、前記代表的な人物検出領域ごとに、前記推定モデルの学習を行い、

前記属性推定部は、前記カメラが生成した画像に写っている人物の検出位置に応じて、前記推定モデルを選択して、前記人物の属性を推定する

ことを特徴とする請求項 8 に記載の人物属性推定システム。

【請求項 10】

属性推定対象人物を撮影して画像を生成するカメラと、前記カメラが生成した画像に写っている人物の属性を推定する属性推定部とを備えた人物属性推定システムにおいて使用される、人物の属性推定のための推定モデルの学習に用いる学習用データを生成する学習用データ生成装置であって、

前記カメラによる前記属性推定対象人物の撮影環境を示す、撮影環境データを取得する撮影環境データ取得部と、

人物画像である標準画像を取得する標準画像取得部と、

前記撮影環境データに応じて前記標準画像のデータを加工することにより、前記標準画像に前記撮影環境を反映させた疑似現場画像を生成する疑似現場画像生成部とを備え、

前記疑似現場画像、又は前記カメラにより前記属性推定対象の人物が撮影される環境において前記カメラで前記疑似現場画像が現出した被写体を撮影することにより得られた疑

10

20

30

40

50

似現場画像撮影画像を用いて前記学習用データを生成することを特徴とする学習用データ生成装置。

【請求項 11】

前記学習用データは、学習用画像データである前記疑似現場画像又は前記疑似現場画像撮影画像と、学習用正解属性データであって、前記疑似現場画像の生成に用いられた前記標準画像に対応付けられた属性データと、を含むことを特徴とする請求項 10 に記載の学習用データ生成装置。

【発明の詳細な説明】

【技術分野】

【0001】

10

本発明は、人物の画像から人物の属性（例えば、年齢や性別）を推定する人物属性推定システム、及び学習用データ生成装置に関し、特に、学習により生成される推定モデルを用いて人物の属性を推定する人物属性推定システム、及び学習用データ生成装置に関するものである。

【背景技術】

【0002】

従来、コンビニエンスストア等の店舗において、来店客の客層を分析するために、店舗の入口付近等において来店客を撮影し、その画像から来店客の属性を推定する人物属性推定システムが知られている。かかる人物属性推定システムは、撮影された映像の各フレーム画像から被写体の顔領域を検出して、検出された顔について、年齢、性別等の属性を推定するものであり、予め作成された属性推定のためのモデルが、製造・出荷の時点で記憶されている。

20

【0003】

例えば、特許文献 1 には、オフライントレーニングシステムを構成するコンピュータで、属性が既知である人物の顔画像をそれぞれ含む複数枚のサンプル画像データと、個々の人物の属性とを対応付けた学習サンプルデータを用いて学習を行うことにより、属性識別辞書を作成するとともに、この属性識別辞書を予め格納するコンピュータにおいて、属性識別辞書を参照して、接続されたカメラにより撮影された人物の属性を識別する技術が開示されている。

【先行技術文献】

30

【特許文献】

【0004】

【特許文献 1】特開 2006 - 323507 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、特許文献 1 のような従来技術においては、学習用のサンプル画像は、属性推定対象となる人物が撮影される環境、すなわち、実際のカメラの設置場所の環境とは異なる環境において撮影されたものであるため、かかるサンプル画像を用いて生成された推定モデルには実際の利用現場の撮影環境が反映されず、したがって、実際の利用に供された際に、精度よく人物の属性を推定することが困難であるという課題があった。

40

【0006】

より具体的には、複数の学習用サンプル画像は、例えば、全て同じ照度の下、全て正面から人物の顔を撮影した画像であるのが一般的である。したがって、当該照度と同じ照度の下で、正面から来店客等を撮影するのであれば、これらのサンプル画像をもとに生成された推定モデルに基づいても、精度よく人物の属性を推定することができるのに対し、当該照度とは異なる照度の下、正面以外の方向から来店客等を撮影する場合には、属性推定の精度が低下することがある。

【0007】

このような精度の低下を防ぐには、店舗等の現場に設置されたカメラにより実際に撮影

50

された画像を学習用サンプル画像として持ち帰り、正解属性データと対応付けた上で改めて推定モデルを生成し直す手法も有効ではある。ところが、かかる手法では、実際の利用現場において、様々な属性の被写体を何人も撮影したり、数千ないし数万というオーダーの大量のサンプル画像の1つ1つに対して、推定モデルを生成するための正解属性データの付与を行うなどの煩雑な作業が必要となる。

【0008】

本発明は、上記の問題に鑑みてなされたものであり、煩雑な作業を要することなく、属性推定の対象となる人物が撮影される環境に応じて、精度よく人物の属性を推定することができる人物属性推定システム、及び学習用データ生成装置を提供することを目的とする。

10

【課題を解決するための手段】

【0009】

本発明の人物属性推定システムは、属性推定対象人物を撮影して画像を生成するカメラと、前記カメラが生成した画像に写っている人物の属性を、推定モデルを用いて推定する属性推定部と、前記カメラによる前記属性推定対象人物の撮影環境を示す、撮影環境データを取得する撮影環境データ取得部と、人物画像である標準画像を取得する標準画像取得部と、前記撮影環境データに応じて前記標準画像のデータを加工することにより、前記標準画像に前記撮影環境を反映させた疑似現場画像を生成する疑似現場画像生成部と、前記疑似現場画像を用いて前記推定モデルの学習を行なう学習部とを備えた構成を有している。

20

【0010】

この構成によれば、どのような撮影環境の下で属性推定対象となる人物が撮影されるかを示す撮影環境データと、人物画像である標準画像とをもとに、実際に人物を現場にて撮影したかのような疑似現場画像が生成され、この疑似現場画像を用いて推定モデルの学習が行われるので、カメラの利用現場の状況やカメラの状態等の撮影環境を反映した、精度の良い属性推定のためのモデルを生成して利用することができる。

【0011】

また、上記の人物属性推定システムにおいて、前記学習部は、前記カメラによる前記属性推定対象人物の撮影環境において、前記カメラで前記疑似現場画像が現出した被写体を撮影することにより得られた疑似現場画像撮影画像を学習用画像データとして用いて前記推定モデルの学習を行うものであってよい。

30

【0012】

この構成によれば、標準画像をもとに生成された疑似現場画像が現出した被写体を、属性推定対象の人物の撮影を行う利用現場又はそれを想定した環境において撮影することで、この撮影画像が、属性推定モデル生成のための学習用サンプル画像として用いられるので、カメラノイズ等、実際の撮影環境をさらによく反映した、精度の良い属性推定のためのモデルを生成して利用することができる。

【0013】

また、上記の人物属性推定システムにおいて、前記学習部は、前記疑似現場画像生成部にて生成された前記疑似現場画像を学習用画像データとして用いて前記推定モデルの学習を行なうものであってよい。

40

【0014】

この構成によれば、標準画像をもとに生成された疑似現場画像が属性推定モデル生成のための学習用サンプル画像として用いられるので、実際の撮影環境を反映した、精度の良い属性推定のためのモデルを、簡易に生成して利用することができる。

【0015】

また、上記の人物属性推定システムにおいて、前記学習部は、前記疑似現場画像の生成に用いられた前記標準画像に対応する属性データを、学習用正解属性データとして用いて前記推定モデルの学習を行うものであってよい。

【0016】

50

この構成によれば、標準画像の属性データを、学習用画像データである疑似現場画像又は疑似現場画像撮影画像の正解属性データとして用いて、推定モデルの学習を行うので、学習用画像データと正解属性データとの対応付けを簡易な構成により実現することができ、属性推定モデルの学習を行うことができる。

【0017】

また、上記の人物属性推定システムは、前記カメラで生成された画像に基づき、前記撮影環境データを算出する撮影環境推定部をさらに備え、前記撮影環境データ取得部は、前記撮影環境推定部で算出された前記撮影環境データを取得するものであってよい。

【0018】

この構成によれば、実際の撮影画像に基づき算出された撮影環境データが、推定モデルの学習に用いられる疑似現場画像の生成に用いられるので、例えば、利用現場において撮影環境データ算出のための画像を撮影することで、撮影画像から把握される撮影環境を、疑似現場画像及び推定モデルに反映させることができる。

10

【0019】

また、上記の人物属性推定システムにおいて、前記撮影環境データは、前記カメラにより前記属性推定対象人物が撮影される場所における照明状態を示すデータを含み、前記疑似現場画像生成部は、前記照明状態を示すデータに応じて、前記標準画像を変換することにより、前記疑似現場画像を生成してよい。

【0020】

この構成によれば、実際の利用現場の照明状態に応じて標準画像を変換し、疑似現場画像が生成されるので、属性推定の精度に影響を与える要因である、利用現場の照明状態を反映した属性推定モデルを生成し、利用することができる。

20

【0021】

また、上記の人物属性推定システムにおいて、前記属性推定部は、前記カメラが生成した画像に写っている人物の属性を、前記画像中の顔領域の部分画像に基づき推定するものであり、前記撮影環境データは、前記属性推定対象人物が、前記カメラにより撮影される場合の顔の向きに関するデータを含み、前記標準画像は、人物の顔を含む画像であって、前記疑似現場画像生成部は、前記顔の向きに関するデータに応じて、前記標準画像の顔の向きを変換することにより、前記疑似現場画像を生成してよい。

【0022】

この構成によれば、属性推定対象となる人物の顔向きとして予想される、撮影環境データとしての顔の向きに関するデータに応じて標準画像を変換し、疑似現場画像が生成されるので、撮影画像の人物の顔領域に着目して行う属性推定に適した推定モデルを生成し、利用することができる。

30

【0023】

また、上記の人物属性推定システムにおいて、前記撮影環境データは、前記カメラにより生成される画像内の1つ以上の代表的な人物検出領域ごとの撮影環境データであってよい。

【0024】

この構成によれば、代表的な人物検出領域、すなわち、撮影画像内で属性推定対象人物の検出が予想される部分領域についての撮影環境データを用いて疑似現場画像が生成されるので、実際の属性推定に適した形で撮影環境データを反映させた推定モデルを生成し、利用することができる。

40

【0025】

また、上記の人物属性推定システムにおいて、前記疑似現場画像生成部は、前記代表的な人物検出領域ごとの撮影環境データを用いて、前記代表的な人物検出領域ごとの疑似現場画像を生成し、前記学習部は、前記代表的な人物検出領域ごとに、前記推定モデルの学習を行い、前記属性推定部は、前記カメラが生成した画像に写っている人物の検出位置に応じて、前記推定モデルを選択して、前記人物の属性を推定するものであってよい。

【0026】

50

この構成によれば、属性推定対象人物の検出が予想される代表的な物検出領域が複数ある場合、各領域ごとに、撮影環境データの反映が異なりうる、代表的な人物検出領域ごとの撮影環境データを用いて、それぞれの疑似現場画像及び推定モデルを生成し、人物属性推定の際には、対象人物がどの代表的な人物検出領域（またはそれに近い位置）に写っているかに応じて、推定モデルを使い分け、より精度よく属性推定を行うことができる。

【0027】

本発明の学習用データ生成装置は、属性推定対象人物を撮影して画像を生成するカメラと、前記カメラが生成した画像に写っている人物の属性を推定する属性推定部とを備えた人物属性推定システムにおいて使用される、人物の属性推定のための推定モデルの学習に用いる学習用データを生成する学習用データ生成装置であって、前記カメラによる前記属性推定対象人物の撮影環境を示す、撮影環境データを取得する撮影環境データ取得部と、人物画像である標準画像を取得する標準画像取得部と、前記撮影環境データに応じて前記標準画像のデータを加工することにより、前記標準画像に前記撮影環境を反映させた疑似現場画像を生成する疑似現場画像生成部とを備え、前記疑似現場画像、又は前記カメラにより前記属性推定対象の人物が撮影される環境において前記カメラで前記疑似現場画像が現出した被写体を撮影することにより得られた疑似現場画像撮影画像を用いて前記学習用データを生成する構成を有している。

【0028】

この構成によれば、どのような撮影環境の下で属性推定対象となる人物が撮影されるかを示す撮影環境データと、標準画像とをもとに、実際に人物を現場にて撮影したかのような疑似現場画像が生成され、この疑似現場画像又は疑似現場画像を利用現場環境の下で撮影した疑似現場画像撮影画像を用いて、属性推定モデルの学習用の画像データが生成されるので、カメラの利用現場の状況やカメラの状態等を反映した、精度の良いモデルの生成のための学習用データを生成することができる。

【0029】

また、上記の学習用データ生成装置において、前記学習用データは、学習用画像データである前記疑似現場画像又は前記疑似現場画像撮影画像と、学習用正解属性データであって、前記疑似現場画像の生成に用いられた前記標準画像に対応する属性データとを含むものであってよい。

【0030】

この構成によれば、疑似現場画像又は疑似現場画像撮影画像を学習用画像データとして、疑似現場画像の生成に用いられた標準画像に対応する属性データが学習用正解データとなるので、簡易に学習用データを生成することができる。

【発明の効果】

【0031】

本発明によれば、撮影環境データに応じて標準画像のデータを加工することにより疑似現場画像が生成され、この疑似現場画像を用いて、属性推定のための推定モデルの学習による生成が行われるので、実際の撮影環境に応じた、精度の良い属性推定が可能となる。

【図面の簡単な説明】

【0032】

【図1】本発明の第1の実施の形態における人物属性推定システムの構成を示すブロック図

【図2】本発明の第1の実施の形態における人物属性システムの動作フロー図

【図3】(a)本発明の第1の実施の形態におけるカメラの設置状況の一例を示す図 (b)本発明の第1の実施の形態におけるカメラで撮影された画像の一例を示す図

【図4】(a)本発明の第1の実施の形態における撮影画像の顔向き角度のデータの分布の一例を示す図 (b)本発明の第1の実施の形態における撮影画像の輝度コントラストのデータの分布の一例を示す図

【図5】(a)本発明の第1の実施の形態における標準画像の一例を示す図 (b)本発

10

20

30

40

50

明の第 1 の実施の形態における標準画像の一例を示す図 (c) 本発明の第 1 の実施の形態における疑似現場画像の一例を示す図 (d) 本発明の第 1 の実施の形態における疑似現場画像の一例を示す図

【図 6】本発明の第 1 の実施の形態における属性推定部の構成を示すブロック図

【図 7】本発明の第 1 の実施の形態における属性推定のための特徴量の変換を説明する図

【図 8】本発明の第 1 の実施の形態における推定モデル再学習部の動作フロー図

【図 9】本発明の第 1 の実施の形態における撮影環境推定部の構成を示すブロック図

【図 10】本発明の第 2 の実施の形態における人物属性推定システムの構成を示すブロック図

【図 11】本発明の第 2 の実施の形態における疑似現場画像撮影画像と正解属性データの関連付けの一例を説明する図 10

【図 12】本発明の第 2 の実施の形態における撮影環境推定部の構成を示すブロック図

【発明を実施するための形態】

【0033】

以下、本発明を実施するための形態について、図面を参照しながら説明する。本発明は、人物の属性を推定する人物属性推定システム、及び学習用データ生成装置に関するが、以下の実施の形態では、人物の属性として、人物の年齢（年代）及び性別を推定する場合について説明する。

【0034】

[第 1 の実施の形態]

20

図 1 は、本発明の第 1 の実施の形態の人物属性推定システムの構成を示す図である。人物属性推定システム 1 は、カメラ 10、再学習制御システム 20、人物属性推定装置 30 を備えている。カメラ 10、再学習制御システム 20、人物属性推定装置 30 は、それぞれ、図示しない通信ユニットを備え、相互に接続される。なお、再学習制御システム 20 は、ネットワーク上のサーバないしサーバ群により実現され、人物属性推定装置 30 とともに、クラウド・コンピューティングシステムを形成する。

【0035】

カメラ 10 は、属性推定の対象となる人物を撮影する。人物属性推定装置 30 は、カメラ 10 で撮影された人物の属性を、顔領域の画像に基づいて推定する。再学習制御システム 20 は、人物属性推定装置 30 が、カメラ 10 で撮影された人物の属性を推定する際に 30

【0036】

カメラ 10 は、年代及び性別を推定したい不特定多数の人物（属性推定対象の人物）が通る場所を撮影するように設置される。例えば、店舗の入口から入ってくる客の顔を撮影するように店舗内の高所に設置される。したがって、属性推定対象となる人物の顔は、カメラ 10 により正面から撮影されるとは限らないし、撮影時の照明や自然光の状態（光源の方向、照度等）などの照明状態も、カメラ 10 が設置される場所や、撮影時刻等により変わりうる。すなわち、カメラ 10 が設置される利用現場が異なれば、属性推定の対象となる人物の画像が、どのような環境で撮影されるかも異なることになる。

【0037】

40

カメラ 10 は、画像生成部 11 と、撮影環境推定部 12 とを備える。このうち、画像生成部 11 は、映像として、連続的なフレーム画像を生成する。あるいは、所定の時間間隔での撮影のたびに、静止画像を生成するものであってもよい。このように生成された画像は、実際に人物属性の推定を行う場合には、人物属性推定装置 30 に提供され、人物属性推定のためのモデルを生成する際には、撮影環境推定部 12 に出力される。

【0038】

撮影環境推定部 12 は、カメラ 10 による人物の撮影環境を推定し、撮影環境を示すデータを算出する。ここで、撮影環境とは、撮影される人物の顔の向き、撮影された顔画像の輝度コントラスト等、生成される画像の内容に影響を与えうる、カメラの設置位置（カメラの設置高さ、レンズ方向等）や、撮影時のカメラ周辺の照明や自然光の状態（光源の 50

方向、照度等)等の、カメラ10自体ないしカメラ10周辺の状態をいう。このような撮影環境は、撮影された画像に反映されうるので、本実施の形態では、撮影された画像をもとに、撮影環境を推定させるデータを算出する。撮影環境を推定させるデータとは、本実施の形態においては、具体的には、生成された画像における人物の顔向き角度(上下・左右)分布と、顔領域における輝度コントラスト分布についてのデータである。算出されたデータは、撮影環境データとして、再学習制御システム20に提供される。

【0039】

再学習制御システム20は、撮影環境データ記憶部21、標準画像記憶部22、疑似現場画像生成部23、再学習用データ記憶部24、推定モデル再学習部25を備える。これらすべてを一つのサーバが備えてもよいし、ネットワーク上で接続される複数のサーバが

10

【0040】

撮影環境データ記憶部21は、カメラ10の撮影環境推定部にて算出された、撮影環境を示すデータである、撮影環境データを記憶する。標準画像記憶部22には、複数の顔画像が、それぞれ、正解値としての属性値と対応付けて記憶されている。本実施の形態では、推定対象となる属性値は、年代と性別であるから、正解値として対応付けられる属性値も、年代(年齢)及び性別を含む。標準画像記憶部22に記憶される標準画像は、異なる利用現場に設置されるカメラ10に共通に提供され得る。したがって、標準画像として記憶される複数の顔画像は、カメラ10の設置場所以外の実験室等で、人物の正面から撮影されたものであってよい。

20

【0041】

疑似現場画像生成部23は、撮影環境データ記憶部21及び標準画像記憶部22からデータを読み出し、撮影環境データを用いて、標準画像データから疑似現場画像を生成する。疑似現場画像とは、標準画像にカメラ10の撮影環境を反映させて生成される画像である。このようにして生成される画像は、利用現場、すなわち、カメラ10が実際に設置される場所において取得されるであろう仮想的な撮影画像とみることができる点で、「疑似現場画像」と呼ぶことができる。生成された疑似現場画像は、再学習用データ記憶部24に出力される。

【0042】

再学習用データ記憶部24には、生成された疑似現場画像と、当該疑似現場画像の生成に用いられた標準画像データに対応付けられた正解値としての属性値とが対応付けられて、再学習用データとして格納されている。また、本実施の形態では、再学習により更新されることとなる、属性推定装置30が予め備える推定モデルに関するデータも格納されている。再学習用データ記憶部24に記憶されるこれらのデータは、推定モデル再学習部25に出力される。

30

【0043】

推定モデル再学習部25は、再学習用データを用いて、推定モデルの再学習を行う。再学習により更新された推定モデルは、人物属性推定装置30に提供される。

【0044】

人物属性推定装置30は、推定モデル記憶部31、属性推定部32を備える。推定モデル記憶部31には、人物属性推定装置30の製造・出荷の時点において、標準画像記憶部22に記憶されるような、汎用的な正解付き顔画像による学習により生成された初期推定モデルが格納されている。推定モデル記憶部31は、推定モデル再学習部25から取得した新たな推定モデルを、属性推定部32に出力する。

40

【0045】

属性推定部32は、カメラ10により撮影された画像を取得し、推定モデルを用いて、当該画像に含まれる人物の年代、性別を推定する。

【0046】

次に、人物属性推定システム1の動作を、図2から図6を用いて説明する。

【0047】

50

まず、実際の利用現場、すなわち、属性推定対象となる人物の撮影場所において、カメラ 10 により撮影された人物の顔を含む画像を取得する（ステップ S 2 1）。ステップ S 2 1 における人物撮影画像取得の目的は、カメラ 10 の撮影環境に関する情報を得ることである。したがって、カメラ 10 を実際の利用現場に設置し、あるいは、実際の利用現場を想定した場所において撮影が行われることが好ましい。

【0048】

カメラ 10 が、映像として、連続的なフレーム画像を生成する場合、同一人物の顔の位置が徐々に変化する複数の連続的なフレーム画像が生成されるが、本実施の形態では、その中から、代表的な顔検出領域に顔が写っているフレーム画像が選択される。代表的な顔検出領域とは、撮影された画像において属性推定対象人物の顔領域の検出が予想される部分領域のことであり、具体的には、カメラ 10 により生成されたフレーム画像において、人物の顔領域が検出される回数が多かった 1 以上の部分領域のことをいう。画像中、どの位置に人物の顔が写るかによって、撮影環境の反映のされ方は異なりうる。実際の利用の際に最も多く顔が検出されると予想される部分領域における撮影環境を、後述のように推定モデルに反映することで、より精度よく属性推定を行うことが可能となる。

【0049】

例えば、図 3（a）に示すように、店舗の入口 D 付近にカメラ 10 を設置して、入口 D 付近を撮影すると、図 3（b）のような画像が得られる。この位置において、D から店舗に入って来る来店客を撮影する場合、図 3（b）内の領域 A 1、A 2 で、最も多く来店客の顔が写ることが分かれば、これらの領域が代表的な顔検出領域となり、領域 A 1 に来店客の顔が写っているフレーム画像と、領域 A 2 に来店客の顔が写っているフレーム画像が選択されることになる。領域 A 1 と領域 A 2 における顔の向き、照明状況等は異なりうるが、これらの領域における顔の向き、照明状況等がほぼ同じであれば、いずれかの領域に顔を含むフレームのみが選択されてもよい。

【0050】

次に、ステップ S 2 1 で撮影された画像を用いて、カメラ 10 の撮影環境を推定させるデータとしての、画像における人物の顔向き角度（上下・左右）と、顔領域における輝度コントラストが算出される（ステップ S 2 2）。本実施の形態では、これらの撮影環境データが分布により表現されるため、ステップ S 2 2 においては、多くの撮影画像を用いることが好ましい。なお、ステップ S 2 2 における撮影環境データの算出処理については、後に詳述する。

【0051】

図 4（a）は、ステップ S 2 2 で算出された、顔向き角度データの分布の一例を模式的にグラフ化したもの、図 4（b）は、顔領域における輝度コントラストデータの分布の一例を模式的にグラフ化したものである。図 4（a）の例では、図 3（b）の領域 A 1 における顔の向きは、下向き 10 度、右向き 20 度を中心として分布しているのに対し、図 3（b）の領域 A 2 における顔の向きは、下向き 20 度、左右 0 度を中心として分布していることが分かる。一方、図 4（b）の例では、領域 A 1 におけるコントラストは、60% が最も多いのに対し、領域 A 2 におけるコントラストは、20% が最も多いことが分かる。

【0052】

次に、ステップ S 2 2 で求めた撮影環境データと、標準画像とを用いて、疑似現場画像が生成される（ステップ S 2 3）。疑似現場画像の生成は、言い換えれば、カメラ 10 の撮影環境を考慮した、標準画像の変換等の加工である。標準画像の加工に、カメラ 10 の撮影環境データが用いられるために、生成された疑似現場画像は、標準画像の被写体である人物を利用現場にて直接撮影して得た画像として利用することができるという意味で、疑似的な現場画像ということができる。疑似現場画像の生成は、ステップ S 2 2 において、撮影環境データの分布を求めたことにあわせて、本実施の形態では、データの分布に割合に応じて行われる。

【0053】

図5(a)は、標準画像記憶部22に記憶される、標準画像の例である。標準画像は、正面から撮影された顔画像である。各画像の被写体である人物の年齢、性別は既知であり、図5(a)、(b)に示すように、各画像に、それらの属性値が正解値として付与されている。この標準画像の顔の向きとコントラストを、ステップS22で算出した撮影環境データに基づき変換する。図4(a)の例では、領域A2における顔の向きは、下向き20度、右向き0度を中心として分布していた。そこで、標準画像の顔の向きを下向き20度、右向き0度及びその付近の角度に変換した画像が、図4(a)の分布割合にしたがって、領域A2における疑似現場画像として生成される。領域A1についても同様である。

【0054】

標準画像の顔の向きの変換は、AAM(Active Appearance Model)や、3D Morphable Model等の種々の方法により行うことが可能である。これらの手法によるモデルを構築する際、上下方向及び左右方向の顔向き変化に対応するパラメータを設定することで、任意の角度に顔の向きを変換することができる。このように、ステップS23では、標準画像の顔向きが3次元的に変換されて疑似現場画像が生成される。

【0055】

さらに、図4(b)の例では、領域A2におけるコントラストは、20%を中心として分布していた。そこで、標準画像のコントラストが20%、及びその付近の値となるように、図4(b)の分布割合にしたがって、変換画像が生成される。領域A1についても同様である。このように標準画像を変換することにより生成された疑似現場画像の例を図5(c)、(d)に示す。図5(c)は、図5(a)の標準画像の変換により、領域A1における疑似現場画像として生成されたもの、図5(d)は、図5(b)の標準画像の変換により領域A2における疑似現場画像として生成されたものである。

【0056】

続いて、属性推定のための推定モデルの再学習が行われる(ステップS24)。本実施の形態においては、属性推定装置30に、予め推定モデルが記憶されており、再学習により、初期推定モデルが更新され、新たな推定モデルが生成される。ステップS23で生成された疑似現場画像は、この推定モデルの再学習に用いられる。より具体的には、生成された疑似現場画像に、正解属性値が付与され、この正解付き疑似現場画像が学習のためのサンプルとして入力される。なお、推定モデルの再学習を行うための処理の流れについては、後に詳述する。

【0057】

疑似現場画像は、各標準画像を変換して生成されるので、疑似現場画像に対応付けられる正解属性値も、元の標準画像に付されたものが用いられる。学習用の新たなサンプル画像の追加に伴い、通常は、その正解属性値を求めることと、それを新たなサンプル画像に対応付けることが必要となる。これに対し、本実施の形態では、新たなサンプル画像である疑似現場画像の被写体は、標準画像の被写体と同一であるため、標準画像の正解属性値をそのまま対応付ければ足り、煩雑な作業は必要ない。

【0058】

図5の例では、図5(c)の疑似現場画像は、図5(a)の標準画像から、図5(d)の疑似現場画像は、図5(b)の標準画像からそれぞれ生成されたものである。したがって、図5(c)の疑似現場画像と、性別が女、かつ、年代が40代という属性値が、また、図5(d)の疑似現場画像と、性別が男、かつ、年代が30代という属性値がそれぞれ組み合わせられ、再学習用データとして用いられる。この例では、撮影画像内の領域A1、A2について、それぞれ異なる疑似現場画像が生成されるため、推定モデルの再学習も、このような部分領域ごとに行い、部分領域ごとに異なる推定モデルが生成されてもよい。

【0059】

そして、このような再学習により得られた推定モデルを用いて、属性推定対象人物の属性推定が行われる(ステップS25)。人物の属性の推定については、後に詳述するが、概略、以下の処理が行われる。すなわち、まず、カメラ10で撮影された来店客の画像が

ら、顔検出を行い、その顔画像から、特徴量の抽出を行う。この特徴量から、ステップS24にて学習により更新した推定モデルを用いて、当該顔画像の人物の属性である、年代、性別を推定する。

【0060】

図5の例において、撮影画像内の複数の部分領域において、異なる推定モデルが生成された場合、カメラ10により撮影された人物の属性推定も、当該人物の顔が撮影画像においてどこに位置するかに応じて、異なる推定モデルを使い分けて行われてもよい。例えば、カメラ10により撮影された画像中、属性推定対象人物の顔が図3(b)の部分領域A1に近い位置にて検出された場合、領域A1について生成された推定モデルを用いて属性推定が行われてよい。また、部分領域A1に近い位置、A2に近い位置の両方で連続的に人物の顔が検出された場合、より鮮明に人物の顔が写っている方の部分領域の推定モデルを用いて、属性推定が行われるようにしてもよい。

10

【0061】

このように、本実施の形態においては、カメラ10の利用現場環境、すなわち、カメラ10が実際に設置される場所及びカメラ10の実際の設置状況を示すデータを用いて、汎用的な顔画像である標準画像が、あたかも、カメラ10の実際の設置場所にて撮影されたものであるかのように変換される。そして、この疑似現場画像を用いて推定モデルに再学習が施され、更新された推定モデルに基づき、属性推定が行われる。したがって、煩雑な作業なしに、利用現場の撮影環境を反映した、属性推定のための推定モデルを生成することができる。

20

【0062】

(属性推定部の詳細な構成)

次に、本実施の形態における属性推定部32の詳細な構成について、図面を参照して説明する。図6は、属性推定部32の構成を示すブロック図である。図6に示すように、属性推定部32は、撮影画像取得部321、顔検出部322、顔特徴抽出部323、属性算出部324を備える。

【0063】

撮影画像取得部321は、撮影により、カメラ10の画像生成部11にて生成された画像を取得し、顔検出部322に出力する。顔検出部322は、撮影画像中の顔領域を検出し、顔領域の部分画像を顔特徴抽出部323に出力する。顔領域の検出は、種々の手法を用いて行うことが可能であり、例えば、Haar-like特徴に基づくAdaBoost法によることができる。なお、前述のように、本実施の形態では、代表的な顔検出領域における撮影環境データが算出され、推定モデルの生成が行われる。したがって、顔検出部322は、代表的な顔検出領域ないしその付近の位置で顔が検出された画像の、顔領域部分画像を顔特徴抽出部323に出力してもよい。

30

【0064】

顔特徴抽出部323は、顔領域の部分画像から、ガボール特徴等、既存の顔特徴量を抽出し、属性算出部324に出力する。本実施の形態では、顔特徴量の抽出の精度向上のため、取得した顔領域画像から、目、鼻、などの顔部品を検出し、これらの顔部品を基準として顔の大きさ等について正規化を行ったうえで、顔特徴量の抽出を行う。

40

【0065】

属性算出部324は、顔特徴抽出部323から取得した顔特徴量から、推定モデル記憶部31に記憶される推定モデルを用いて、顔領域の部分画像の属性を求める。本実施の形態では、線形判別法に従い、取得した一般的な顔特徴量を、属性推定のための特徴量空間に射影した上で、属性推定関数を用いて、属性値を算出する。つまり、本実施の形態において、推定モデルとは、顔特徴抽出部323から取得した顔特徴量を、属性推定のための特徴量空間に射影する行列と、射影後の特徴量空間で属性推定を行うための属性推定関数のことを指す。

【0066】

線形判別法では、属性推定のための特徴量空間への射影を、以下の式(1)によって行

50

う。

【数 1】

$$y = W^T x \quad \text{式 (1)}$$

ここで、 x は、射影前の顔特徴量ベクトル、 y は射影後の特徴量ベクトルである。また、 W は、写像行列であり、以下では、特徴空間射影行列と呼ぶ。

【0 0 6 7】

図 7 は、特徴空間射影行列 W により、既存の手法により抽出された顔特徴量 x が、属性推定のための特徴量 y に変換された例を示す図である。線形判別法により、変換後の顔特徴量の次元は、変換前よりも小さくなっているが、推定対象属性（年代、性別）をより適切に表現する特徴量となる。

10

【0 0 6 8】

一方、属性推定関数は、射影後の特徴量ベクトル y を入力値とする関数であり、以下の式 (2) により求められる。

【数 2】

$$f(y) = b^T \cdot y \quad \text{式 (2)}$$

ここで、

20

【数 3】

$$b = (Y^T Y + \alpha I)^{-1} Y t \quad \text{式 (3)}$$

であり、 α は、重み係数、 I は単位行列である。 t は、各正解属性値 t_i を要素とするベクトルあり、各 t_i は、例えば、年代であれば、20（代）、30（代）等、性別なら、-1（男性）、+1（女性）等の値をとる。ベクトル t は、サンプルが k 個ある場合、以下のように記述することができる。

【数 4】

$$t = \begin{bmatrix} t_1 \\ \vdots \\ t_k \end{bmatrix} \quad \text{式 (4)}$$

30

また、

【数 5】

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1d} \\ \vdots & \ddots & \vdots \\ y_{k1} & \cdots & y_{kd} \end{bmatrix} \quad \text{式 (5)}$$

40

であり、 d は射影後の特徴量ベクトル y の次元数である。

式 (2) に射影後の特徴量ベクトルを入力すると、その顔画像の人物の属性値を表すスカラー量が出力される。

【0 0 6 9】

（推定モデルの再学習の処理の流れ）

次に、本実施の形態における、推定モデル再学習部 25 における推定モデル再学習の処理の流れを、図 8 のフロー図を用いて処理の流れを説明する。

50

【 0 0 7 0 】

まず、再学習用データが取得される（ステップ S 8 1）。再学習用データは、上述のように、疑似現場画像とそれに対応する正解属性値データ、及び、再学習により更新されるべき初期推定モデルに関するデータである。次に、ステップ S 8 1 で取得された各疑似現場画像の顔特徴量を抽出する（ステップ S 8 2）。ここでの顔特徴量は、ガボール特徴量等の既存の顔特徴量でよい。

【 0 0 7 1 】

そして、ステップ S 8 2 で抽出された、疑似現場画像の顔特徴量と、ステップ S 8 1 で取得された各疑似現場画像に対応する正解属性値及び初期推定モデルに関するデータとを用いて、特徴空間射影行列 W の再学習による更新が行われる（ステップ S 8 3）。 10

【 0 0 7 2 】

本実施の形態の属性推定において用いられる線形判別法では、特徴空間射影行列 W は、グループ間分散の、グループ内分散に対する比が最大となるような特徴空間を生成する行列として定義される。したがって、ステップ S 8 3 における特徴空間射影行列 W の再学習も、グループ間分散の、グループ内分散に対する比がより大きくなるよう、以下の式（6）を解くことにより行われる。

【数 6】

$$W = \arg \max \frac{|W^T \Sigma_B' W|}{|W^T \Sigma_W' W|} \quad \text{式 (6)}$$

20

【 0 0 7 3 】

ここで、 Σ_B は、属性グループ間の共分散行列、 Σ_W は、属性グループ内の共分散行列である。属性グループ間共分散行列 Σ_B は、ある属性について、異なる属性値のグループ同士の相関傾向を示す指標として機能する。これに対して、属性グループ内共分散行列 Σ_W は、ある属性について、同じ属性値のグループ内において、個々のデータの相関傾向を示す指標として機能する。例えば、属性として性別に着目すると、属性グループ間共分散行列 Σ_B が、男性のグループと女性のグループとがどのくらい離れているかを示すのに対し、属性グループ内共分散行列 Σ_W は、男性 / 女性それぞれのグループ内で、個々の特徴量データのばらつきの程度がどのくらいであることを示す。 30

【 0 0 7 4 】

式（6）における属性グループ間共分散行列 Σ_B' 、属性グループ内共分散行列 Σ_W' は、ステップ S 8 1 で取得した正解属性値とステップ S 8 2 で抽出した疑似現場画像の顔特徴量データを用いて更新（再計算）された、属性グループ間共分散行列 Σ_B 、属性グループ内共分散行列 Σ_W である。このように、ステップ S 8 3 においては、2 つの共分散行列がまず更新される。

【 0 0 7 5 】

属性グループ間共分散行列 Σ_B 、属性グループ内共分散行列 Σ_W は、それぞれ、以下の式（7）、式（8）により、 Σ_B' 、 Σ_W' に更新される。 40

【数 7】

$$\Sigma_B' = (1 - \lambda) \Sigma_{B_retrain} + \lambda \Sigma_{B_ini} \quad \text{式 (7)}$$

【数 8】

$$\Sigma_W' = (1 - \lambda) \Sigma_{W_retrain} + \lambda \Sigma_{W_ini} \quad \text{式 (8)}$$

ここで、 $\Sigma_{B_retrain}$ 及び $\Sigma_{W_retrain}$ は、再学習用の疑似現場画像についての、属性グループ間共分散行列と、属性グループ内共分散行列であり、ステップ S 8 1 で取得した正 50

解属性値とステップ S 8 2 で抽出した疑似現場画像の顔特徴量データを用いて算出される。

【 0 0 7 6 】

また、 B_{ini} 及び W_{ini} は、初期共分散行列であり、疑似現場画像を用いた再学習前の属性グループ間共分散行列と、属性グループ内共分散行列である。ステップ S 8 1 で取得した、初期推定モデルに関するデータとは、この初期共分散行列のことである。本実施の形態において、これらの初期共分散行列は、例えば、疑似現場画像を生成するのに用いられた正解属性値付きの標準画像を用いて、予め作成されていてよい。また、 α は 0 から 1 の間の値をとる重み係数である。例えば、再学習用のデータが十分蓄積されていれば、 α をより大きく設定して、共分散行列、ひいては、特徴空間射影行列 W に、疑似現場画像の特徴量をより反映させることができる。

10

【 0 0 7 7 】

属性グループ間共分散行列 B 、属性グループ内共分散行列 W は、それぞれ、以下の式、式 (9)、式 (1 0) から求められる。

【数 9】

$$\Sigma_B = \sum_{j=1}^C (\mu_j - \mu)(\mu_j - \mu)^T \quad \text{式 (9)}$$

【数 1 0】

20

$$\Sigma_W = \sum_{j=1}^C \sum_{i=1}^{n_j} (x_{ji} - \mu_j)(x_{ji} - \mu_j)^T \quad \text{式 (1 0)}$$

このうち、 C は、属性グループ数である。本実施の形態においては、性別については $C = 2$ 、年代については $C = 10$ (0 から 10 歳、10 から 20 歳、... 90 から 100 歳) とする。また、 n_j は属性グループ番号 j のサンプル数、 μ_j は属性グループ j の平均顔特徴量、 μ は全サンプルの平均顔特徴量、 x_i は個々の画像の顔特徴量を、それぞれ示す。

【 0 0 7 8 】

以上のようにして、特徴空間射影行列 W を再学習により更新したのち、属性推定関数の更新が行われる (ステップ S 8 4)。属性推定関数は、上述のように、特徴空間射影後の特徴量 y の関数 $f(y)$ として記述することができるが、射影前の特徴量ベクトル x と、特徴空間射影行列 W とを用いて、式 (1 1) のように、 x の関数 $f(x)$ としても記述することができる。

30

【数 1 1】

$$f(x) = b^T \cdot y = b^T W^T x \quad \text{式 (1 1)}$$

したがって、特徴空間射影行列 W の更新に伴い、属性推定関数 $f(x)$ も更新されることとなる。

40

【 0 0 7 9 】

以上のように、本実施の形態では、属性推定のためのモデルである特徴空間射影行列 W 及び属性推定関数 $f(x)$ が、疑似現場画像の特徴量と、正解属性値を用いた再学習により更新される。

【 0 0 8 0 】

(撮影環境推定部の詳細な構成)

次に、本実施の形態における撮影環境推定部 1 2 の、詳細な構成について、図面を参照しながら説明する。図 9 は、撮影環境推定部 1 2 の構成を示すブロック図である。図 9 に示すように、撮影環境推定部 1 2 は、入力画像取得部 1 2 1、顔検出部 1 2 2、顔向き推定部 1 2 3、照明状態推定部 1 2 4、統計部 1 2 5 を備える。

50

【 0 0 8 1 】

入力画像取得部 1 2 1 は、撮影された画像、より具体的には、カメラ 1 0 の画像生成部 1 1 で生成されたフレーム画像を取得し、顔検出部 1 2 2 に出力する。顔検出部 1 2 2 は、取得された画像内の顔領域を検出し、取得された画像のうち、顔領域の部分画像を、画像内の位置情報とともに、顔向き推定部 1 2 3、照明状態推定部 1 2 4、統計部 1 2 5 に出力する。

【 0 0 8 2 】

顔向き推定部 1 2 3 は、顔検出部 1 2 2 から取得した部分画像に含まれる顔の向きを推定し、顔向きに関するデータを統計部 1 2 5 に出力する。顔向き角度の推定は、種々の手法により行うことが可能である。例えば、学習用サンプル顔画像の特徴量と、顔検出部 1 2 2 から取得した顔画像の特徴量との距離に基づき、k 近傍法を用いて推定することができる。

10

【 0 0 8 3 】

照明状態変動部 1 2 4 は、取得した顔領域の部分画像における輝度コントラストを算出し、統計部 1 2 5 に出力する。輝度コントラスト C は、以下の式 (1 2) により算出される。

【 数 1 2 】

$$C = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad \text{式 (1 2)}$$

20

式 (1 2) において、 I_{\min} は、顔領域の最小輝度値、 I_{\max} は、顔領域の最大輝度値である。

【 0 0 8 4 】

統計部 1 2 5 は、入力画像である各フレーム画像内の顔検出領域についてクラスタリングを行い、最も多く顔が検出される画像内の領域を、代表的な顔検出領域として特定する。また、統計部 1 2 5 は、顔向き推定部 1 2 3 から取得したデータを用いて、代表的な顔検出領域における顔の向きの分布を算出する。統計部 1 2 5 は、さらに、照明状態推定部 1 2 4 から取得したデータを用いて、代表的な顔検出領域における輝度コントラストの分布を算出する。統計部 1 2 5 は、顔向き分布データと、輝度コントラストデータを、撮影環境データ記憶部 2 1 に出力する。

30

【 0 0 8 5 】

このように、撮影環境推定部 1 2 は、撮影によりカメラ 1 0 内で生成された画像をもとに、カメラ 1 0 による撮影環境に関するデータを算出するための構成を備えている。

【 0 0 8 6 】

以上のように、本発明の第 1 の実施の形態の人物属性推定システムによれば、標準画像をカメラ 1 0 の撮影環境を示すデータを用いて変換した画像である疑似現場画像を再学習用のサンプル画像とし、標準画像に対応付けられた属性データを、疑似現場画像の正解属性データとして用いて、属性推定のためのモデルの再学習を行うので、煩雑な作業を要することなく、カメラ 1 0 の実際の利用現場の撮影環境を反映したモデルを再構築することができ、年代、性別といった人物の属性の推定の精度を向上できる。

40

【 0 0 8 7 】

なお、上記の実施形態では、撮影環境データとして、顔向きと照明状態に関するデータを用いたが、このほかに、イメージセンサノイズや J P E G 圧縮等によるブロック歪み、ピントの合い方等、カメラ自体の性質又はカメラの設定に起因する、種々のカメラノイズに関するデータを用いてもよい。カメラノイズに関するデータは、画像生成部 1 0 から取得した、利用現場の撮影画像を用いて算出してもよいし、カメラノイズのに関するデータが既知であれば、それを入力してもよい。

【 0 0 8 8 】

また、撮影環境データとして、カメラ 1 0 周辺の照度の時間的な変動を加味し、時間帯

50

ごとに異なる疑似現場画像と推定モデルを生成し、属性推定の時刻に応じて推定モデルを選択し、人物属性推定を行ってもよい。

【 0 0 8 9 】

[第 2 の実施形態]

次に、本発明の第 2 の実施の形態の人物属性推定システムを説明する。第 2 の実施形態において、標準画像からカメラの利用現場の撮影環境を反映した疑似現場画像を生成する点は、第 1 の実施形態と同様である。第 2 の実施形態では、さらに推定精度を向上させるべく、生成された疑似現場画像に、カメラ 1 0 の実際の利用現場でのカメラノイズ等を反映させた新たな画像を生成し、これを再学習用の画像データとするための構成を備えている。

10

【 0 0 9 0 】

図 1 0 は、第 2 の実施の形態の人物属性システムの構成を示す図である。カメラ 1 0 と属性推定装置 3 0 の構成は、第 1 の実施の形態と同じである。また、再学習制御システム 2 0 において、撮影環境データ記憶部 2 1、標準画像記憶部 2 2、疑似現場画像生成部 2 3、再学習用データ記憶部 2 4、推定モデル再学習部 2 5 を備える点も、第 1 の実施形態と同じである。これに加え、本実施の形態では、再学習制御システム 2 0 は、さらに、疑似現場画像記憶部 2 6、疑似現場画像出力部 2 7、疑似現場画像撮影画像取得部 2 8 を備える。

【 0 0 9 1 】

疑似現場画像記憶部 2 6 は、疑似現場画像生成部 2 2 において生成された、疑似現場画像を記憶する。疑似現場画像出力部 2 7 は、疑似現場画像記憶部 2 6 に格納される疑似現場画像を出力する。疑似現場画像出力部 2 7 は、プリンタやタブレット P C 等、図示しない所定のユニットに接続され、これらのユニットに対して、疑似現場画像データを出力する。出力された疑似現場画像は、例えば、紙等への印刷や、タブレット P C のディスプレイ上に表示されるなどして、可視化される。

20

【 0 0 9 2 】

カメラ 1 0 の画像生成部 1 1 は、疑似現場画像出力部 2 7 から出力され、疑似現場画像が現出された被写体（印刷物やディスプレイなど）がカメラ 1 0 で撮影された場合、その画像を生成し、再学習制御システム 2 0 に出力する。

【 0 0 9 3 】

疑似現場画像撮影画像取得部 2 8 は、画像生成部 1 1 で生成された、疑似現場画像の撮影画像としての画像を取得し、再学習用データ記憶部 2 4 に出力する。すなわち、疑似現場画像撮影画像とは、疑似現場画像が現出した被写体をカメラ 1 0 の利用現場において、カメラ 1 0 で撮影した画像である。このように、疑似現場画像を利用現場で実際に撮影することで、カメラのノイズや、照明状況を、よりリアルに再現した画像を得ることができる。

30

【 0 0 9 4 】

再学習用データ記憶部 2 4 は、第 1 の実施の形態と同様、更新対象となる推定モデルに関するデータとして、初期共分散行列データを記憶する。また、再学習用データ記憶部 2 4 は、再学習用画像データとして、疑似現場画像撮影画像を記憶し、再学習用正解属性データとして、標準画像に対応付けられた正解属性値を記憶する。

40

【 0 0 9 5 】

再学習用画像データである疑似現場画像撮影画像と、再学習用正解属性データである標準画像の属性値との関連付けは、種々の方法により実現される。例えば、図 1 1 に示すように、疑似現場画像出力部 2 7 が、標準画像の正解属性値を用いて、属性値（ここでは属性として性別に着目）ごとに疑似現場画像を並べて出力するとともに、属性が切り替わるところで、顔画像でないことが識別可能なマーカとなるデータを挿入して出力する構成としてもよい。出力された順序で撮影を行うことで、疑似現場画像撮影画像と正解属性値データとの対応付けを、属性値ごとにまとめて簡易に行うことができる。これに代えて、疑似現場画像出力部 2 7 が、各疑似現場画像に正解属性値や対応する標準画像の I D を示す

50

バーコードとともに疑似現場画像を出力し、撮影されたバーコードに基づき、疑似現場画像撮影画像に属性値の関連付けを行ってもよい。

【0096】

推定モデル再学習部25は、再学習用データ記憶部24に記憶されるデータを用いて、属性推定モデルの再学習を行う。つまり、本実施の形態においては、疑似現場画像の代わりに、疑似現場画像撮影画像を用いて、その特徴量を抽出した上、推定モデルの再学習を行う。

【0097】

図12は、第2の実施の形態における撮影環境推定部12の構成を示す図である。本実施の形態のように、疑似現場画像を実際にカメラ10の利用現場にて撮影する場合、カメラのノイズと、当該利用現場における照明状態は、その疑似現場画像が現出した被写体を撮影した画像から知ることができる。したがって、本実施の形態においては、撮影環境推定部12は、第1の実施の形態における照明状態推定部124を備えていない。

10

【0098】

本実施の形態においても、第1の実施の形態におけるのと同様、統計部125は、代表的な顔検出領域を特定する。カメラ10による疑似現場画像の撮影は、統計部125により特定された代表的な顔検出領域に、疑似現場画像の顔がくるような位置に、疑似現場画像の可視化物を提示して行うことが好ましい。

【0099】

以上のように、第2の実施の形態の属性推定システム1によれば、カメラ10の撮影環境を反映した疑似現場画像を生成し、この疑似現場画像が現出した被写体をカメラ10で撮影した画像を用いて推定モデルの再学習を行うので、カメラ10の撮影環境をより正確に反映した推定モデルを生成することができる。

20

【0100】

なお、疑似現場画像から3次元の人物モデルを生成して、これを被写体とし、疑似現場画像撮影画像を生成してもよい。

【0101】

[その他の変形例]

上記の実施の形態では、カメラ10の生成画像から被写体の顔領域を検出し、被写体の属性を推定する場合について説明したが、顔領域以外の人物の部分領域画像を用いて、属性を推定してもよい。この場合、上記の実施形態における代表的な顔検出領域に替えて、代表的な人物検出領域における撮影環境データを取得して、疑似現場画像を生成し、推定モデルを生成してもよい。さらに、推定対象となる属性は、年齢(年代)又は性別のみであってもよいし、年齢、性別に限られず、人種や社会的地位ないし区分(高校生、社会人、老人等)であってもよい。

30

【0102】

また、上記の実施の形態では、撮影環境データとして、顔の向きに関するデータ及び照明状態に関するデータ、カメラノイズを例に挙げて説明したが、撮影環境データは、撮影画像に影響を与える他の要因に関するものであってもよいし、上記の実施の形態において、いくつかの撮影環境データを任意に組み合わせて、又は、いずれか1つの撮影環境データのみを用いて、疑似現場画像を生成してもよい。

40

【0103】

また、上記の実施の形態では、人物の画像である標準画像を加工することで疑似現場画像を生成する場合について説明したが、例えば、ステレオカメラにより人物を撮影し、人物の3次元画像を生成するためのポリゴンデータのような、人物モデルデータを加工して疑似現場画像を生成してもよい。

【0104】

また、上記の実施の形態では、撮影環境推定部11をカメラ10が備え、推定モデル記憶部31を属性推定装置30が備える場合について説明したが、撮影環境推定部及び/または推定モデル記憶部とを再学習制御システム20が備えていてもよい。また、再学習制

50

御サーバの構成を、カメラ１０ないし属性推定装置３０が備えていてもよいし、カメラ１０が属性推定装置３０の構成を備えていてもよい。さらに、カメラ１０、再学習制御サーバ２０、属性推定装置３０の動作が全て、１つの装置により実現されてもよい。

【０１０５】

また、上記の実施の形態では、属性推定装置３０が、初期推定モデルを備え、初期推定モデルが、疑似現場画像を用いた再学習により更新される場合について説明したが、属性推定装置３０は、初期推定モデルを備えていなくてもよく、疑似現場画像を用いた学習により生成された推定モデルを初めから属性推定に用いてもよい。また、推定モデルの再学習は、カメラの設置場所の変更や、標準画像の追加等に伴い、繰り返されてよい。

【０１０６】

また、上記の実施の形態では、撮影環境データを、カメラ１０により生成された撮影画像から算出する場合について説明したが、撮影環境データは、ユーザにより手動で入力されてもよいし、カメラ１０の利用現場に設置されたセンサ等から取得されてもよい。例えば、カメラ１０が、店舗に入って来る客の顔をその正面から左に３０度、上に２０度傾いた方向から撮影するように設置されている場合には、撮影画像において、来店客の顔は右に約３０度、下に約２０度向いた状態で撮影されることになる。撮影される顔向きは、人物の身長によって変わりうるので、各年代の平均身長をこれに加味してもよい。このように、撮影環境を直接的に示すデータを得ることができる場合、ステップＳ１における撮影画像の取得ないしその前提となるカメラ１０による撮影はされなくてもよい。

【０１０７】

また、上記の実施の形態では、線形判別法に従い、属性推定する場合について説明したが、カーネル回帰やガウス混合分布モデルなどを用いて属性を推定してもよい。上記の実施の形態においては、初期推定モデルに関するデータが、予め再学習用データ記憶部に記憶されていたが、推定モデルの再学習を行う際に、推定モデル記憶部から初期推定モデル又はそれに関するデータを取得して再学習が行われてもよい。

【産業上の利用可能性】

【０１０８】

本発明の属性推定システムは、実際の撮影環境に応じて、精度の良い属性推定を行うことができるという効果を有し、学習により生成される推定モデルを用いて人物の属性を推定する人物属性推定システム等として有用である。

【符号の説明】

【０１０９】

- １ 人物属性推定システム
- １０ カメラ
- １１ 画像生成部
- １２ 撮影環境推定部
- １２１ 入力画像取得部
- １２２ 顔検出部
- １２３ 顔向き推定部
- １２４ 照明状態推定部
- １２５ 統計部
- ２０ 再学習制御システム
- ２１ 撮影環境データ記憶部
- ２２ 標準画像記憶部
- ２３ 疑似現場画像生成部
- ２４ 再学習用データ記憶部
- ２５ 推定モデル再学習部
- ２６ 疑似現場画像記憶部
- ２７ 疑似現場画像出力部
- ２８ 疑似現場画像撮影画像取得部

10

20

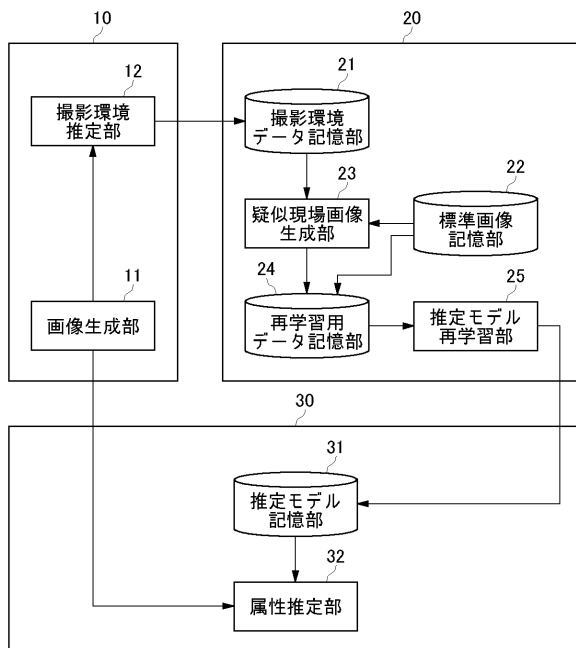
30

40

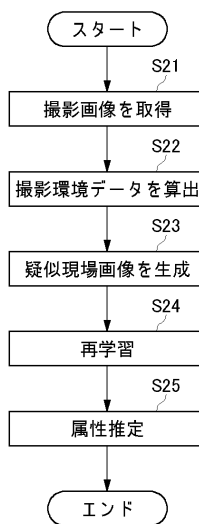
50

- 3 0 人物属性推定装置
- 3 1 推定モデル記憶部
- 3 2 属性推定部
- 3 2 1 撮影画像取得部
- 3 2 2 顔検出部
- 3 2 3 顔特徴抽出部
- 3 2 4 属性算出部

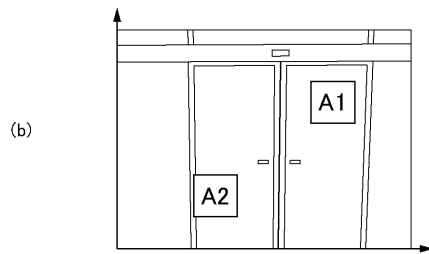
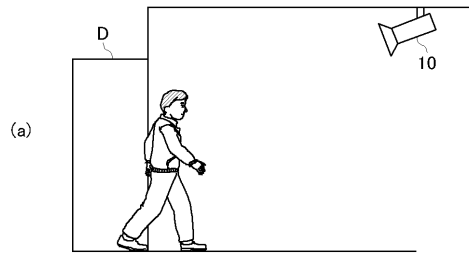
【図 1】



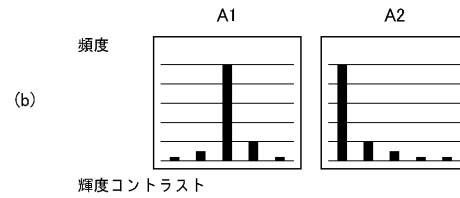
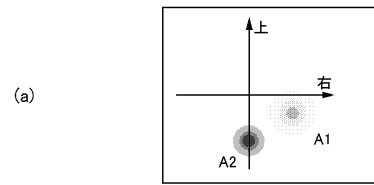
【図 2】



【図 3】



【図 4】



【図 5】



性別	女
年齢	32

(a)



性別	男
年齢	41

(b)

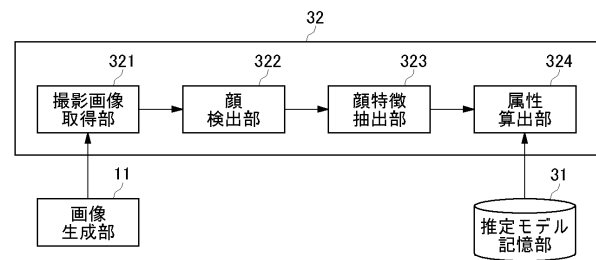


(c)

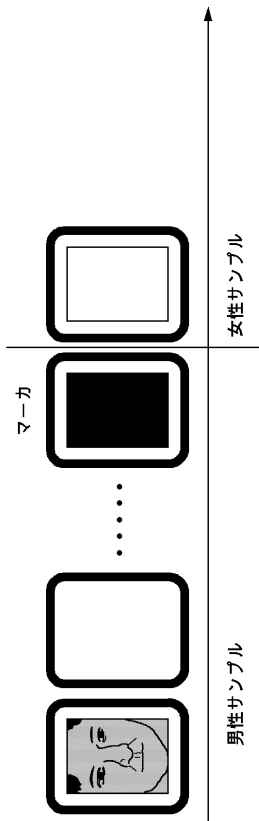


(d)

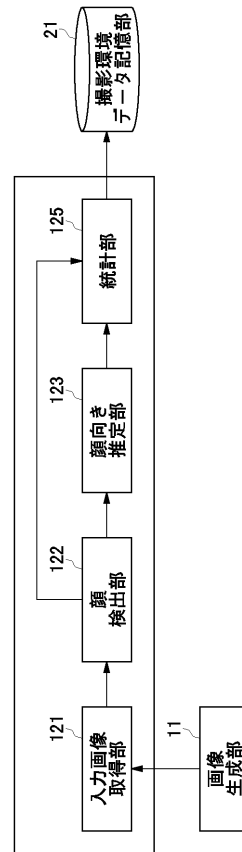
【図 6】



【図 1 1】



【図 1 2】



フロントページの続き

- (72)発明者 由雄 宏明
福岡県福岡市博多区美野島4丁目1番62号 パナソニックシステムネットワークス株式会社内
- (72)発明者 山田 伸
福岡県福岡市博多区美野島4丁目1番62号 パナソニックシステムネットワークス株式会社内
- (72)発明者 松川 隆行
福岡県福岡市博多区美野島4丁目1番62号 パナソニックシステムネットワークス株式会社内

審査官 佐田 宏史

- (56)参考文献 特開2011-141740(JP,A)
特開2011-070471(JP,A)
特開2010-218051(JP,A)
特開2006-031103(JP,A)
特開2005-227794(JP,A)
西山 正志、山口 修, “物体表面の見え属性分類に基づくSelf Quotient Imageを用いた顔認識”, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2005年10月20日, Vol.105, No.374, pp.33-38
島野 美保子、外4名, “任意照明下顔認識のための顔表面の位置相関を考慮したMAP推定”, 情報処理学会論文誌, 日本, 社団法人情報処理学会, 2006年 7月15日, Vol.47, No.SIG 10(CVIM 15), pp.162-172
中野渡 祥裕、外3名, “向き・照明変動にロバストな顔認証”, パナソニック技報, 日本, パナソニック株式会社, 2009年 1月15日, Vol.54, No.4, pp.24-29

(58)調査した分野(Int.Cl., DB名)

G06T 1/00, 7/00 - 7/60
H04N 7/18
G08B 13/194 - 13/196