

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2018-116589

(P2018-116589A)

(43) 公開日 平成30年7月26日(2018.7.26)

(51) Int.Cl.		F I			テーマコード (参考)
<b>G06T 7/00</b>	<b>(2017.01)</b>	G06T 7/00	350B	5B057	
<b>G06T 1/00</b>	<b>(2006.01)</b>	G06T 1/00	340A	5L096	
<b>G06T 7/20</b>	<b>(2017.01)</b>	G06T 7/20	300B		

審査請求 未請求 請求項の数 17 O L (全 23 頁)

(21) 出願番号 特願2017-8044 (P2017-8044)  
 (22) 出願日 平成29年1月20日 (2017.1.20)

(71) 出願人 000208891  
 K D D I 株式会社  
 東京都新宿区西新宿二丁目3番2号  
 (74) 代理人 100135068  
 弁理士 早原 茂樹  
 (74) 代理人 100141313  
 弁理士 辰巳 富彦  
 (72) 発明者 呉 剣明  
 埼玉県ふじみ野市大原二丁目1番15号  
 株式会社K D D I 総合研究所内  
 (72) 発明者 湯 津津  
 埼玉県ふじみ野市大原二丁目1番15号  
 株式会社K D D I 総合研究所内

最終頁に続く

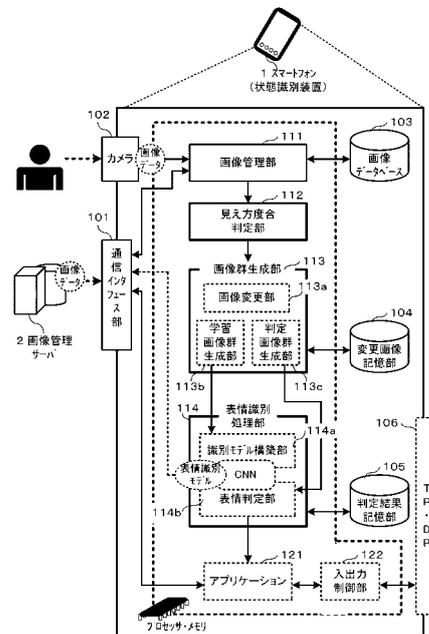
(54) 【発明の名称】 対象画像の変更画像群を用いる状態識別装置、プログラム及び方法

(57) 【要約】

【課題】 様々な状況下で種々の見え方をしている識別対象の画像であっても、より高い精度でこの識別対象の状態を識別することが可能な装置を提供する。

【解決手段】 識別対象に係る画像を用い、識別モデルによって識別対象の状態を識別する本状態識別装置は、入力された識別対象に係る1つの画像に対し、この1つの画像における識別対象の見え方を変更し且つその状態を維持する処理を施し、この処理で生成された画像を含む画像群を生成する画像群生成手段と、識別モデルを構築する際、及び/又は識別モデルを用いて状態を識別する際に、生成された画像群を使用する識別処理手段とを有する。ここで、画像群生成手段は、この1つの画像に対し、見え方に係る少なくとも1つの種別について、この1つの画像における方向とは反対の方向への変更を少なくとも行うことも好ましい。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する状態識別装置であって、

入力された当該識別対象に係る 1 つの画像に対し、当該 1 つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成する画像群生成手段と、

当該識別モデルを構築する際、及び / 又は当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用する識別処理手段とを有することを特徴とする状態識別装置。

10

**【請求項 2】**

前記画像群生成手段は、当該 1 つの画像に対し、当該見え方に係る少なくとも 1 つの種別について、当該 1 つの画像における方向とは反対の方向への変更を少なくとも行うことを特徴とする請求項 1 に記載の状態識別装置。

**【請求項 3】**

前記画像群生成手段は、当該 1 つの画像に対し、当該見え方に係る少なくとも 1 つの種別について、該少なくとも 1 つの種別に関し予め設定された方向であって識別に好適な方向への変更を少なくとも行うことを特徴とする請求項 1 又は 2 に記載の状態識別装置。

**【請求項 4】**

前記画像群生成手段は、当該 1 つの画像に対し、当該見え方に係る複数の種別のうちの複数について合わせて変更を行うことを特徴とする請求項 1 から 3 のいずれか 1 項に記載の状態識別装置。

20

**【請求項 5】**

前記画像群生成手段は、当該 1 つの画像に対し、当該見え方に係る少なくとも 1 つの種別について、該少なくとも 1 つの種別に関し予め設定された方向への変更を行い、当該方向によって特徴付けられる当該識別モデルを構築するための画像群を生成し、

前記識別処理手段は、生成された当該画像群を使用して当該識別モデルを構築することを特徴とする請求項 1 から 4 のいずれか 1 項に記載の状態識別装置。

**【請求項 6】**

前記画像群生成手段は、当該 1 つの画像に対し、当該見え方に係る少なくとも 1 つの種別について、該少なくとも 1 つの種別に関しとり得る全ての方向への変更を行い、該少なくとも 1 つの種別についての方向毎の識別結果のばらつきが抑制された当該識別モデルを構築するための画像群を生成し、

30

前記識別処理手段は、生成された当該画像群を使用して当該識別モデルを構築することを特徴とする請求項 1 から 4 のいずれか 1 項に記載の状態識別装置。

**【請求項 7】**

前記画像群生成手段は、正解として 1 つの状態が対応付けられた 1 つの画像に対し、当該見え方に係る少なくとも 1 つの種別について、当該 1 つの状態に対し予め設定された変更数だけの変更を行い、所定のサンプル数の当該画像群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 1 項に記載の状態識別装置。

40

**【請求項 8】**

前記画像群生成手段は、当該 1 つの状態が対応付けられた画像のサンプル数が、他の状態が対応付けられた画像のサンプル数よりも小さい場合、当該 1 つの状態が対応付けられた画像について、当該他の状態が対応付けられた画像についてよりも大きい変更数だけの変更を行うことを特徴とする請求項 7 に記載の状態識別装置。

**【請求項 9】**

前記識別処理手段は、当該識別モデルを構築する際、及び当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用し、

前記画像群生成手段は、当該状態の識別の際に使用される画像群を生成する場合、入力された 1 つの画像に対し、当該識別対象の見え方に係るいずれの種別についても、当該識

50

別モデル構築の際に使用された画像群を生成する際に行われた当該種別についての変更と同じ方向への変更を行い、一方、該画像群を生成する際に当該種別についての変更が行われなかった場合には変更を行わない

ことを特徴とする請求項 1 から 8 のいずれか 1 項に記載の状態識別装置。

【請求項 10】

前記識別処理手段は、当該識別モデルを用いて、生成された当該画像群に含まれる画像の各々について当該状態を決定し、当該画像が変更前の元画像か否かに応じて予め設定された重み、及び / 又は当該画像における変更された種別に応じた予め設定された重みによって、決定された当該状態の重み付け平均を行い、当該重み付け平均結果を当該状態の識別結果とする

10

ことを特徴とする請求項 1 から 9 のいずれか 1 項に記載の状態識別装置。

【請求項 11】

前記画像群生成手段は、当該識別対象に係る 1 つの画像を含む当該画像群を生成することを特徴とする請求項 1 から 10 のいずれか 1 項に記載の状態識別装置。

【請求項 12】

前記状態識別装置は、入力された当該識別対象に係る 1 つの画像に対し、当該識別対象に係る画像領域を抽出し、当該画像領域に基づいて当該見え方に係る少なくとも 1 つの種別における当該種別発生の有無及び / 又は度合いを判定する度合い判定手段を更に有し、

前記画像群生成手段は、当該少なくとも 1 つの種別について、判定された当該有無の状態及び / 又は当該度合いを変えることによって変更を行う

20

ことを特徴とする請求項 1 から 11 のいずれか 1 項に記載の状態識別装置。

【請求項 13】

前記画像群生成手段は、当該見え方に係る複数の種別として、当該 1 つの画像における当該識別対象の表れている度合い、当該識別対象の表れている向き若しくは傾き、当該識別対象を遮蔽するものの有無若しくは種類、当該識別対象の明度、当該識別対象の色合い、及び当該識別対象と背景とのコントラストの度合いのうち少なくとも 1 つについて変更を行うことを特徴とする請求項 1 から 12 のいずれか 1 項に記載の状態識別装置。

【請求項 14】

当該識別対象は顔であって、当該状態は顔の表情であり、

前記画像群生成手段は、入力された当該識別対象に係る 1 つの画像に対し、当該 1 つの画像における顔の見え方を変更し且つ該顔の表情を維持する処理を施し、

30

前記識別処理手段は、当該識別モデルを用いて、入力された 1 つの画像における顔の表情を識別する

ことを特徴とする請求項 1 から 13 のいずれか 1 項に記載の状態識別装置。

【請求項 15】

識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する状態識別システムであって、

入力された当該識別対象に係る 1 つの画像に対し、当該 1 つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成する画像群生成手段と、

40

生成された当該画像群によって当該識別モデルを構築する識別モデル構築手段と、

生成された当該画像群を使用し、当該識別モデルによって当該状態を識別する識別処理手段と

を有することを特徴とする状態識別システム。

【請求項 16】

識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する装置に搭載されたコンピュータを機能させる状態識別プログラムであって、

入力された当該識別対象に係る 1 つの画像に対し、当該 1 つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成する画像群生成手段と、

50

当該識別モデルを構築する際、及び/又は当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用する識別処理手段としてコンピュータを機能させることを特徴とする状態識別プログラム。

【請求項17】

識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する装置に搭載されたコンピュータにおいて実施される状態識別方法であって、

入力された当該識別対象に係る1つの画像に対し、当該1つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成するステップと、

当該識別モデルを構築する際、及び/又は当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用するステップとを有することを特徴とする状態識別方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、所定対象の状態を、当該所定対象に係る画像に基づいて識別する技術に関する。

【背景技術】

【0002】

従来、所定対象の状態、例えば人間の表情を、この所定対象に関する画像、例えば顔を撮影した写真画像を用いて識別する技術は、種々考案されてきた。

【0003】

特に、人間の表情認識の分野では、ポジティブ、ネガティブ、ニュートラルの3分類モデルや、Paul Ekmanの7分類モデル(ニュートラル、喜び、嫌悪、怒り、サプライズ、悲しみ、恐怖)等を用いて、多くの研究者が表情認識技術の向上に取り組んでいる。

【0004】

このような取り組みの一例として、特許文献1には、上記の分類モデルに基づく大量の顔画像データの特徴量を学習し、その特徴量に基づいて表情を識別する技術が開示されている。この技術では、特に、意図的に作った顔ではなく自然な顔表情の学習データを効率良く収集し、認識精度の良い識別器を作成することを目的としている。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2011-150381号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、特許文献1に記載されたような従来技術では、実際の表情認識環境に耐え得る高い識別精度を達成するために、様々な状況下で撮影された、種々の撮影具合の大量の顔画像が、学習データとして必要となってくる。

【0007】

この学習データが大量に得られず、識別器が現実の様々な状況での種々の撮影具合の顔画像を十分に学習することができない場合、顔の表情を誤って判定するケースも多く発生し、結果として識別精度が低下し、実用に堪えない又は汎用性の低い識別器となってしまう。

【0008】

例えば、真正面を向いた顔の画像における表情識別では、ある一定の識別精度に達している識別器においても、識別すべき顔画像において、顔の一部が画像範囲(画像枠)からはみ出していたり、顔が上下左右のいずれかの方に向いていたたり、頭部の傾きが大きかったり、帽子やサングラス等の部品が付いていたたり、さらには、画像の顔部分の照度(明度)

10

20

30

40

50

が低すぎたり高すぎたりしている場合、識別精度が顕著に低下してしまうことも少なくない。

【0009】

また一方で、このような様々な状況での種々の撮影具合の顔画像を、学習データとして大量に準備することは極めて困難なのが実情となっている。その結果、ある顔画像における表情の識別において、当該顔画像の撮影された状況や撮影の具合によっては、表情の識別精度が低いまま処理を行い、表情を誤って識別してしまう事態が数多く発生してしまう。

【0010】

そこで、本発明は、様々な状況下で種々の見え方をしている識別対象の画像であっても、より高い精度でこの識別対象の状態を識別することが可能な装置、プログラム及び方法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】

【0011】

本発明によれば、識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する状態識別装置であって、

入力された当該識別対象に係る1つの画像に対し、当該1つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成する画像群生成手段と、

当該識別モデルを構築する際、及び/又は当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用する識別処理手段とを有する状態識別装置が提供される。

20

【0012】

この本発明による状態識別装置の一実施形態として、画像群生成手段は、当該1つの画像に対し、当該見え方に係る少なくとも1つの種別について、当該1つの画像における方向とは反対の方向への変更を少なくとも行うことも好ましい。

【0013】

また、本発明による状態識別装置の他の実施形態として、画像群生成手段は、当該1つの画像に対し、当該見え方に係る少なくとも1つの種別について、この少なくとも1つの種別に関し予め設定された方向であって識別に好適な方向への変更を少なくとも行うことも好ましい。

30

【0014】

さらに、本発明による状態識別装置の更なる他の実施形態として、画像群生成手段は、当該1つの画像に対し、当該見え方に係る複数の種別のうちの複数について合わせて変更を行うことも好ましい。

【0015】

また、本発明による状態識別装置の更なる他の実施形態として、画像群生成手段は、当該1つの画像に対し、当該見え方に係る少なくとも1つの種別について、この少なくとも1つの種別に関し予め設定された方向への変更を行い、当該方向によって特徴付けられる当該識別モデルを構築するための画像群を生成し、

40

識別処理手段は、生成された当該画像群を使用して当該識別モデルを構築することも好ましい。

【0016】

さらに、本発明による状態識別装置の更なる他の実施形態として、画像群生成手段は、当該1つの画像に対し、当該見え方に係る少なくとも1つの種別について、この少なくとも1つの種別に関しとり得る全ての方向への変更を行い、この少なくとも1つの種別についての方向毎の識別結果のばらつきが抑制された当該識別モデルを構築するための画像群を生成し、

識別処理手段は、生成された当該画像群を使用して当該識別モデルを構築することも好ましい。

50

## 【0017】

また、本発明による状態識別装置の更なる他の実施形態として、画像群生成手段は、正解として1つの状態が対応付けられた1つの画像に対し、当該見え方に係る少なくとも1つの種別について、当該1つの状態に対し予め設定された変更数だけの変更を行い、所定のサンプル数の当該画像群を生成することも好ましい。

## 【0018】

ここで、上記の変更数設定に係る実施形態において、画像群生成手段は、当該1つの状態が対応付けられた画像のサンプル数が、他の状態が対応付けられた画像のサンプル数よりも小さい場合、当該1つの状態が対応付けられた画像について、当該他の状態が対応付けられた画像についてよりも大きい変更数だけの変更を行うことも好ましい。

10

## 【0019】

さらに、本発明による状態識別装置の更なる他の実施形態として、識別処理手段は、当該識別モデルを構築する際、及び当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用し、

画像群生成手段は、当該状態の識別の際に使用される画像群を生成する場合、入力された1つの画像に対し、当該識別対象の見え方に係るいずれの種別についても、当該識別モデル構築の際に使用された画像群を生成する際に行われた当該種別についての変更と同じ方向への変更を行い、一方、この画像群を生成する際に当該種別についての変更が行われなかった場合には変更を行わない

ことも好ましい。

20

## 【0020】

また、本発明による状態識別装置の更なる他の実施形態として、識別処理手段は、当該識別モデルを用いて、生成された当該画像群に含まれる画像の各々について当該状態を決定し、当該画像が変更前の元画像か否かに応じて予め設定された重み、及び/又は当該画像における変更された種別に応じて予め設定された重みによって、決定された当該状態の重み付け平均を行い、当該重み付け平均結果を当該状態の識別結果とすることも好ましい。

## 【0021】

さらに、本発明による状態識別装置の更なる他の実施形態として、画像群生成手段は、当該識別対象に係る1つの画像を含む当該画像群を生成することも好ましい。

30

## 【0022】

また、本発明による状態識別装置の更なる他の実施形態として、状態識別装置は、入力された当該識別対象に係る1つの画像に対し、当該識別対象に係る画像領域を抽出し、当該画像領域に基づいて当該見え方に係る少なくとも1つの種別における当該種別発生の有無及び/又は度合いを判定する度合い判定手段を更に有し、

画像群生成手段は、当該少なくとも1つの種別について、判定された当該有無の状態及び/又は当該度合いを変えることによって変更を行うことも好ましい。

## 【0023】

さらに、本発明による状態識別装置において、具体的に、画像群生成手段は、当該見え方に係る複数の種別として、当該1つの画像における当該識別対象の表れている度合い、当該識別対象の表れている向き若しくは傾き、当該識別対象を遮蔽するものの有無若しくは種類、当該識別対象の明度、当該識別対象の色合い、及び当該識別対象と背景とのコントラストの度合いのうちの少なくとも1つについて変更を行うことも好ましい。

40

## 【0024】

また、本発明による状態識別装置において、具体的に、当該識別対象は顔であって当該状態は顔の表情であり、

画像群生成手段は、入力された当該識別対象に係る1つの画像に対し、当該1つの画像における顔の見え方を変更し且つこの顔の表情を維持する処理を施し、

識別処理手段は、当該識別モデルを用いて、入力された1つの画像における顔の表情を

50

識別する  
ことも好ましい。

【0025】

本発明によれば、また、識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する状態識別システムであって、

入力された当該識別対象に係る1つの画像に対し、当該1つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成する画像群生成手段と、

生成された当該画像群によって当該識別モデルを構築する識別モデル構築手段と、

生成された当該画像群を使用し、当該識別モデルによって当該状態を識別する識別処理手段と  
を有する状態識別システムが提供される。

10

【0026】

本発明によれば、さらに、識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する装置に搭載されたコンピュータを機能させる状態識別プログラムであって、

入力された当該識別対象に係る1つの画像に対し、当該1つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成する画像群生成手段と、

当該識別モデルを構築する際、及び/又は当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用する識別処理手段としてコンピュータを機能させる状態識別プログラムが提供される。

20

【0027】

本発明によれば、さらにまた、識別対象に係る画像を用い、識別モデルによって当該識別対象の状態を識別する装置に搭載されたコンピュータにおいて実施される状態識別方法であって、

入力された当該識別対象に係る1つの画像に対し、当該1つの画像における当該識別対象の見え方を変更し且つ当該状態を維持する処理を施し、当該処理で生成された画像を含む画像群を生成するステップと、

当該識別モデルを構築する際、及び/又は当該識別モデルを用いて当該状態を識別する際に、生成された当該画像群を使用するステップと  
を有する状態識別方法が提供される。

30

【発明の効果】

【0028】

本発明の状態識別装置、プログラム及び方法によれば、様々な状況下で種々の見え方をしている識別対象の画像であっても、より高い精度でこの識別対象の状態を識別することができる。

【図面の簡単な説明】

【0029】

【図1】本発明による状態識別装置の一実施形態における機能構成を示す機能ブロック図である。

40

【図2】本発明による状態識別方法の一実施形態の概略を示すフローチャートである。

【図3】見え方度合判定部での判定に係る見え方種別の例を説明するための顔画像データである。

【図4】本発明に係る変更画像作成・画像群生成処理の一実施形態を示すフローチャートである。

【図5】識別モデル構築部で構築される識別モデルの一実施形態を示す模式図である。

【図6】本発明による状態識別システムの一実施形態を示す模式図である。

【発明を実施するための形態】

【0030】

50

以下、本発明の実施形態について、図面を用いて詳細に説明する。

【0031】

[状態識別装置]

図1は、本発明による状態識別装置の一実施形態における機能構成を示す機能ブロック図である。

【0032】

図1によれば、本実施形態の状態識別装置としてのスマートフォン1は、公知の構成を有するカメラ102を内蔵しており、このカメラ102を用いて、例えばユーザの顔を撮影してこの顔の写真画像を生成し、生成した写真画像に映ったユーザの顔の表情を識別して、例えばタッチパネル・ディスプレイ(TP・DP)106に識別結果を表示することができる。また、当然に、このような表情識別対象である顔の写真画像を、外部から通信ネットワーク経由で通信インタフェース部101を介して取得し、表情識別処理を行うことも可能である。

10

【0033】

さらに、1つの応用例として、スマートフォン1のアプリケーション121、例えば対話AIAアプリが、この表情の識別結果を利用して、例えば対話しているユーザの感情(発話意図)を理解し、その応答内容を調整したり、当該ユーザとの対話内容をパーソナライズしたりすることも可能になる。

【0034】

さらに、スマートフォン1は、本実施形態において、表情識別のための表情識別処理部114における学習用の(人間の顔の写真)画像を、画像管理サーバ2から通信インタフェース部101を介して取得することも好ましい。この際、スマートフォン1は、画像群生成部113によって、取得した学習用の画像をベースとして学習に供する学習画像群を生成する。その結果、従来ほど大量の画像サンプルを外部から取り込まずとも、十分な量の学習用画像を準備することができるのである。

20

【0035】

このような本発明による状態識別装置としてのスマートフォン1は、識別対象(例えば人間の顔)に係る画像を用い、識別モデルによって識別対象の状態(例えば顔の表情)を識別する装置であって、

(A)入力された識別対象に係る1つの画像に対し、この1つの画像における識別対象(例えば顔)の見え方を変更し且つその状態(例えば表情)を維持する処理を施し、この処理で生成された画像を含む画像群を生成する画像群生成手段(画像群生成部113)と、(B)識別モデルを構築する際、及び/又は識別モデルを用いて状態(例えば表情)を識別する際に、上記(A)で生成された画像群を使用する識別処理手段(表情識別処理部114)と

30

を有することを特徴としている。

【0036】

ここで、識別モデルを構築する際には、実用に堪える識別精度を達成するために、通常、学習データとして大量の識別対象に係る画像が必要となる。しかも、識別対象の見え方について互いに異なっている画像も大量に含まれることが好ましい。しかしながら、現実には、種々の見え方をしている識別対象画像を大量に準備することは極めて困難であった。また、識別モデルを用いて状態を識別する際にも、識別対象画像における識別対象の見え方によっては、識別モデルによる状態識別結果がその影響を受け、状態を誤って識別する場合も少なくないのが実情であった。

40

【0037】

これに対し、上記構成(A)及び(B)を備えたスマートフォン1によれば、識別対象の見え方を変更し且つその状態を維持する処理で生成された画像を含む画像群を生成する。これにより、例えば、種々の見え方をしている識別対象に係るより多数の画像である学習用及び/又は識別用の画像群を、確実に準備することも可能となる。その結果、様々な状況下で種々の見え方をしている識別対象の画像であっても、より高い精度でこの識別対

50

象の状態を識別することができるのである。

【0038】

なお、上記のスマートフォン1に具現された本発明による状態識別装置は、当然に、識別すべき対象の状態が人間の顔の表情である場合にのみ適用されるものではない。本発明によれば、1つの状態を表した画像におけるその見え方によって状態識別結果が影響を受けるような識別対象の状態であっても、より高い精度で識別可能となる。例えば、人間や動物の性別、年齢、人種(種別)や、服装等を識別すべき状態とすることもできる。さらには、家具、日用品や、自動車等の種々の物体認識にも本発明を適用することが可能である。従来、そのような状態識別の結果が、画像での見え方によって大きな誤差を含んでいたり間違ったりしてきたのに対し、本発明によれば、そのような状態をより精度良く識別することができるのである。

10

【0039】

さらに、スマートフォン1に具現された本発明による状態識別装置は、当然にスマートフォンに限定されるものではない。例えば、この状態識別装置として、タブレット型コンピュータ、ノート型コンピュータ、パーソナルコンピュータ(PC)、セットトップボックス(STB)、ロボット、デジタルサイネージ等を採用することもできる。例えば、カメラを内蔵したこれらの装置(端末)において、様々な実環境下においても、ユーザの表情を読み取ることによって、読み取った表情に係る情報に応じた応答を行ったり、読み取った表情に係る情報から先に実施されたユーザに対するアクション等の評価を行ったりすることも可能となる。

20

【0040】

[状態識別方法]

図2は、本発明による状態識別方法の一実施形態の概略を示すフローチャートである。ここで、本実施形態において、識別対象は人間の顔であって、識別対象の状態はこの顔の表情となっている。また、図2(A)は、本発明に係る学習画像群を用いた表情識別モデル構築処理(学習処理)のフローを示しており、図2(B)は、本発明に係る判定画像群を用いた表情識別処理のフローを示している。以下、最初に、図2(A)を用いて表情識別モデル構築処理の概略を説明する。

【0041】

(S101) 学習用画像を取得する。この学習用画像は、この後、学習画像群を生成して学習処理時の画像サンプル数を増加させるための元画像となる。

30

【0042】

(S102) 取得した学習用画像から顔画像領域を抽出する。

(S103) 抽出した顔画像領域において、予め設定された見え方種別毎に、当該見え方種別発生の有無及び/又は度合いを判定する。

【0043】

ここで、見え方種別としては、例えば、顔が正面を向いている1つの状態、及び顔が正面を向いた状態から見て上下左右の各々の方に向いている4つの状態を含む、顔の向きに係る種別を予め設定することができる。また、この見え方種別における度合いとして、正面向きを基準(ゼロ度)として顔が向いている向きのなす角度としてもよい。なお、この見え方種別については、後に図3を用いて詳細に説明する。

40

【0044】

(S104、S105) 取得した学習用画像における、見え方種別毎における当該見え方種別発生の有無及び/又は度合いの判定結果に基づいて、変更画像を作成し、当該変更画像を含む学習画像群を生成する。これらの変更画像作成処理及び学習画像群生成処理については、後に図4を用いて詳細に説明する。

(S106) 生成した学習画像群を用いて表情識別モデルを構築する。これで、表情識別モデル構築処理が完了する。

【0045】

次いで、図2(B)を用いて表情識別処理の概略を説明する。

50

(S201) 識別対象画像を取得する。この識別対象画像は、状態識別の対象であるとともに、この後、判定画像群を生成して状態識別処理時の画像サンプル数を増加させるための元画像となる。

【0046】

(S202) 取得した識別対象画像から顔画像領域を抽出する。

(S203) 抽出した顔画像領域において、予め設定された見え方種別毎に、当該見え方種別発生の有無及び/又は度合いを判定する。この見え方種別については、ステップS103で説明したものと同様とすることができる。いずれにしても、この見え方種別について、後に図3を用いて詳細に説明する。

【0047】

(S204、S205) 取得した識別対象画像における、見え方種別毎における当該見え方種別発生の有無及び/又は度合いの判定結果に基づいて、変更画像を作成し、当該変更画像を含む判定画像群を生成する。これらの変更画像作成処理及び学習画像群生成処理は、ステップS104及びS105での処理と同様のものとしたり、互いに関連付けられた連携した処理としたりすることができる。いずれにしても、これらの処理について、後に図4を用いて詳細に説明する。

【0048】

(S206) 生成した判定画像群の各々に対し、生成された表情識別モデルを用いて、その顔の表情についての判定を行う。ここで使用する表情識別モデルは、ステップS106で構築されたものとする。または、別途、例えば公知の方法で予め生成された表情識別モデルを利用することも可能である。

【0049】

(S207) ステップS206での判定結果を総合して、取得した識別対象画像における顔の表情の識別を行う。これで、表情識別処理が完了する。

【0050】

[装置構成]

次に図1に戻って、以下、本発明による状態識別装置の具体的構成について説明を行う。同図の機能ブロック図に示すように、状態識別装置(表情識別装置)である本実施形態のスマートフォン1は、通信インタフェース部101と、カメラ102と、画像データベース103と、変更画像記憶部104と、判定結果記憶部105と、タッチパネル・ディスプレイ(TP・DP)106と、プロセッサ・メモリとを有する。ここで、プロセッサ・メモリは、スマートフォン1のコンピュータを機能させるプログラムを実行することによって、状態識別機能(表情識別機能)を実現させる。

【0051】

さらに、このプロセッサ・メモリは、機能構成部として、画像管理部111と、見え方度合判定部112と、画像変更部113a、学習画像群生成部113b及び判定画像群生成部113cを含む画像群生成部113と、識別モデル構築部114a及び表情判定部114bを含む表情識別処理部(状態識別処理部)114と、アプリケーション121と、入出力制御部122とを有する。ここで、図1におけるスマートフォン1の機能構成部間を矢印で接続して示した処理の流れは、本発明による状態識別方法(表情識別方法)の一実施形態としても理解される。

【0052】

通信インタフェース部101は、

(a) 学習画像群生成の元画像となる学習用画像や、表情識別処理部114での学習用の大量の画像のデータ、さらには、

(b) 判定画像群生成の元画像となる識別対象画像や、表情識別処理部114でそのまま識別処理を受ける識別対象画像のデータ

を、例えば外部の画像管理サーバ2からインターネット等の通信ネットワークを介して取得し、画像管理部111に出力する。また、この通信インタフェース部101を介して、本発明に係る表情識別プログラム(アプリ)や、当該表情識別結果を利用したサービスを

10

20

30

40

50

提供可能なアプリケーション・プログラム、例えば対話 AI アプリ、をダウンロードすることも可能となっている。

【0053】

カメラ102は、例えばユーザの顔を撮影し、識別対象画像としての顔画像データを画像管理部111に出力することができる。

【0054】

画像管理部111は、通信インタフェース部101やカメラ105から上述したような画像データを入力し、画像データベース103に保存して管理する。また、ユーザ等による指示や装置内処理からの要請等に応じて、画像群生成の元画像となる学習用画像又は識別対象画像を、見え方度合判定部112に出力する。

10

【0055】

同じく図1において、見え方度合判定部112は、入力された識別対象(例えば顔)に係る1つの画像(学習用画像又は識別対象画像)に対し、識別対象(例えば顔)に係る画像領域を抽出し、この画像領域に基づいて、見え方に係る少なくとも1つの種別における当該種別発生の有無及び/又は度合いを判定する。ここで、種別発生有りを(度合い) = 1とし、種別発生無しを(度合い) = 0と解釈することも可能である。ちなみに、この後、画像群生成部113は、当該種別につき、ここで判定された有無の状態及び/又は度合いを変えることによって元画像の変更を行い、変更画像を作成するのである。

【0056】

なお、これらの画像領域の抽出、及び見え方種別に係る判定は、例えば、画像をスキャンし、画素値が急激に変化するエッジを取り出して識別対象(例えば顔)を検知する公知の方法を用いて実施することができる。また、このような抽出を可能とするオープンソースを利用してよい。さらには、公知の深層学習(Deep Learning)等の機械学習を用いた抽出手法を利用することも可能である。

20

【0057】

図3は、見え方度合判定部112での判定に係る見え方種別の例を説明するための顔画像データである。

【0058】

見え方種別の一例として、図3(A)には、顔が画面からはみ出している画像が示されている。このように、見え方種別として例えば、顔が画面からはみ出していない1つの状態、及び顔が画面の上下左右の各々の辺からはみ出している状態を含む、顔のはみ出しに係る種別を設定することができる。ここで、この見え方種別における度合いとしては、例えば、顔全体の(推定)面積又は画素数に対する、画像の外にはみ出した顔部分の面積又は画素数((顔全体面積又は画素数) - (画像内に残っている顔部分の面積又は画素数))の割合としてもよい。

30

【0059】

また、見え方種別の他の例として、図3(B)には、顔の向き・傾きの角度が大きい画像が示されている。このように、見え方種別として例えば、顔が正面を向いている1つの状態、及び顔が正面を向いた状態から見て上下左右の各々の方に向いている状態を含む、顔の向き・傾きに係る種別を設定してもよい。ここで、この見え方種別における度合いとしては、例えば、正面向きを基準(ゼロ度)として顔が向いている向きのなす角度(及び顔の傾き具合を表す傾き角)とすることができる。

40

【0060】

なお、顔の向きに係る見え方種別の判定は、例えば、上下左右の各々の方に向いている顔の画像によって、予めこれら4つの向きのパターンを学習した識別モデルを生成し、この識別モデルを用いて顔の向きを判定する手法を利用することにより実施可能である。また、顔の傾きに係る見え方種別の判定についても、同様に傾き具合の異なる複数のパターンを学習した識別モデルを生成し、この識別モデルを用いて顔の傾きを判定する手法を利用することにより実施可能である。

【0061】

50

さらに、見え方種別の更なる他の例として、図3(C)には、顔又は顔付近にサングラス、帽子やマスクといった装着物が付された画像が示されている。このように、見え方種別として例えば、顔又は顔付近に装着物が何ら付されていない1つの状態、並びに顔又は顔付近に眼鏡、サングラス、ゴーグル、帽子、頬かむり、スカーフ、マスク及びその他の装着物の各々が付された状態を含む、装着物付加に係る種別を設定することができる。ここで、この見え方種別における度合いとしては、例えば、顔全体の(推定)面積又は画素数に対する、装着物によって遮蔽された顔部分の面積又は画素数(顔全体面積又は画素数)-(遮蔽されていない露出した顔部分の面積又は画素数)の割合としてもよい。

#### 【0062】

なお、このような装着物付加に係る見え方種別の判定も、例えば、所定の装着物を付された顔の画像によって、予め当該装着物の有無を学習した識別モデルを生成し、この識別モデルを用いて装着物の有無を判定する手法を利用することにより実施可能である。

10

#### 【0063】

また、見え方種別の更なる他の例として、図3(D)には、顔の照度(明度)が高すぎたり低すぎたりする画像が示されている。このように、見え方種別として例えば、顔画像領域の全画素における明度(照度)の平均が所定の基準範囲内である1つの状態、及び顔画像領域の全画素における明度(照度)の平均が所定の基準範囲を超えて高すぎる又は低すぎる状態を含む、顔の照度(明度)に係る種別を設定してもよい。ここで、この見え方種別における度合いとしては、例えば、この照度(明度)の平均値と、照度(明度)基準値との差とすることができる。

20

#### 【0064】

なお、当然に、見え方度合判定部112での判定に係る見え方種別は、上述した4つに限定されるものではない。そのうちの1つ、2つ又は3つであってもよく、また、他の見え方種別を含んでいてもよい。ここで、他の見え方種別としては、例えば、顔画像領域の全画素における色(彩度)の平均に係る見え方種別、顔画像における顔画像領域と背景領域との明度/彩度のコントラストに係る見え方種別や、顔部分のサイズ(例えば、画像全体に対する顔部分の面積又は画素数の割合)を挙げることができる。

#### 【0065】

すなわち、見え方種別の例をまとめると、1つの画像における(a)識別対象の表れている度合い、(b)識別対象の表れている向き若しくは傾き、(c)識別対象を遮蔽するものの有無若しくは種類、(d)識別対象の明度、(e)識別対象の色合い、及び(f)識別対象と背景とのコントラストの度合い等のうちの少なくとも1つを、見え方種別として採用することができる。このように、画像中において、識別対象が、その識別対象となる状態を維持したままで判別可能な変化を生じ得るならば、その変化に係る一連の見え方を、見え方種別とすることが可能である。

30

#### 【0066】

図1に戻って、画像群生成部113の画像変更部113aは、入力された識別対象(例えば顔)に係る1つの画像(学習用画像又は識別対象画像)に対し、この1つの画像における識別対象(例えば顔)の見え方を変更し且つ当該状態を維持する処理を施す。

#### 【0067】

次いで、画像群生成部113の学習画像群生成部113b及び判定画像群生成部113cはそれぞれ、画像変更部113aでの処理で作成された変更画像を含む学習画像群及び判定画像群を生成する。生成された画像群(及び元画像)は、変更画像記憶部104に保存され、適宜読み出されて使用されることも好ましい。

40

#### 【0068】

ここで、学習画像群に含まれる各画像には、元画像である学習用画像に付与されている「正解情報」がそのまま付加されている。識別対象の状態として顔の表情を識別する場合、この正解情報としては、ポジティブ、ニュートラル、及びネガティブという表情に関する3つのカテゴリを表す値(例えばそれぞれ1、0及び-1)を用いることができる。当然に、他の表情分類モデル、例えばPaul Ekmanの7つのカテゴリ(ニュートラル、喜び、

50

嫌悪、怒り、サプライズ、悲しみ、恐怖)に基づく感情分類モデルや、さらに細分化された感情分類モデル(例えば、Paul Ekmanモデルの7つのカテゴリに対し、さらに、おもしろさ、軽蔑、満足、困惑、興奮、罪悪感、功績に基づく自負心、安心、納得感、喜び、及び恥を追加したモデル)におけるカテゴリを表す値を、正解情報とすることも可能である。

【0069】

なお、生成される画像群には、入力された1つの画像(元画像)を含めることも好ましい。勿論、元画像を含まない画像群を生成してもよいが、元画像を含めた画像群を用いることによって、より識別精度の高い識別モデルを構築したり、識別モデルによる識別精度を向上させたりすることが可能となる。ここで、以上に説明した画像群生成部113における画像変更・画像群生成処理がとり得る種々の実施形態について、以下に説明する。

10

【0070】

(形態a)入力された1つの画像(学習用画像又は識別対象画像)に対し、少なくとも1つの見え方種別について、この1つの画像における方向とは反対の方向への変更を少なくとも行う。

【0071】

この形態aでは、例えば、(図3(B)で説明した)顔の向き・傾きに係る見え方種別について、入力された1つの画像が左向きの顔の画像である場合、この左向きの顔を右向きに変更する(例えば反転させる)処理を行った変更画像を作成し、この変更画像を少なくとも含む画像群を生成することができる。例えば、左向きの顔の画像(元画像)と、右向きの顔の画像(変更画像)とを含む画像群(学習画像群又は判定画像群)を生成してもよい。ちなみにこのような処理によって、構築する識別モデルの汎用性(画像での見え方に対するロバスト性)を向上させたり、識別モデルの偏向性に対応して識別精度を向上させたりすることも可能となる。

20

【0072】

(形態b)入力された1つの画像(学習用画像又は識別対象画像)に対し、少なくとも1つの見え方種別について、当該見え方種別に関し予め設定された方向であって識別に好適な方向への変更を少なくとも行う。

【0073】

この形態bでは、例えば、(図3(D)で説明した)顔の照度(明度)に係る見え方種別について、入力された1つの画像における顔画像領域の全画素での明度(照度)の平均値が所定の基準明度(照度)範囲を超えて高すぎる場合、この明度(照度)平均値がこの基準明度(照度)範囲の値となるように各画素での明度(照度)を変更する処理を行った変更画像を作成し、この変更画像を少なくとも含む画像群を生成することができる。例えば、明度(照度)の平均値が高すぎる画像(元画像)と、明度(照度)の平均値を所定の基準明度(照度)範囲内に収めた画像(変更画像)とを含む画像群(学習画像群又は判定画像群)を生成してもよい。ちなみにこのような処理によって、より識別性能の高い識別モデルを構築したり、識別モデルによる識別結果の精度を向上させたりすることも可能となる。

30

【0074】

(形態c)入力された1つの画像(学習用画像又は識別対象画像)に対し、予め設定された複数の見え方種別のうちの複数について合わせて変更を行う。

40

【0075】

この形態cでは、例えば、(図3(B)で説明した)顔の向き・傾きに係る見え方種別と、(図3(D)で説明した)顔の照度(明度)に係る見え方種別との両種別について、同時に変更の施された変更画像を作成することができる。具体的には、例えば、左向きの顔の画像であって且つ顔画像領域での明度(照度)の平均値が高すぎる画像(元画像)から、右向きの顔の画像であって且つ明度(照度)の平均値が所定の基準明度(照度)範囲内に収まった画像(変更画像)を作成し、画像群(学習画像群又は判定画像群)に含めてもよい。ちなみにこのような処理によって、構築する識別モデルの汎用性(見え方の種々

50

の変化に対するロバスト性)を向上させたり、識別モデルによる識別結果の精度を向上させたりすることも可能となる。

【0076】

(形態d)入力された1つの画像(学習用画像又は識別対象画像)に対し、少なくとも1つの見え方種別について、当該見え方種別に関し予め設定された方向への変更を行い、当該方向によって特徴付けられる識別モデルを構築するための画像群を生成する。

【0077】

この形態dでは、例えば、(図3(B)で説明した)顔の向き・傾きに係る見え方種別について、入力された1つの画像が下向きではない顔の画像である場合、この顔を下向きに変更する処理を行った変更画像を作成し、この変更画像を少なくとも含む画像群を生成することができる。すなわち、この場合、下向きの顔の画像を必ず(又は高い頻度で)含む画像群(学習画像群又は判定画像群)を生成することができる。ちなみにこのような処理によって、所定の方向によって特徴付けられる(例えば、特に下向きの顔の表情を精度良く識別可能な)識別モデルを構築したり、所定の方向によって特徴付けられた識別モデルの当該特徴に合わせた画像入力を行うことにより識別精度を向上させたりすることも可能となる。

10

【0078】

(形態e)入力された1つの画像(学習用画像又は識別対象画像)に対し、少なくとも1つの見え方種別について、当該見え方種別に関しとり得る全ての方向への変更を行い、当該見え方種別についての方向毎の識別結果のばらつきが抑制された識別モデルを構築するための画像群を生成する。

20

【0079】

この形態eでは、例えば、(図3(B)で説明した)顔の向き・傾きに係る見え方種別について、入力された1つの画像が左向きの顔の画像である場合、この左向きの顔を、右向き、上向き、及び下向きに変更する処理を行った変更画像を作成し、元画像とこれらの変更画像とを含む画像群(学習画像群又は判定画像群)を生成することができる。ちなみにこのような処理によって、識別精度のより高い識別モデルを構築し、識別モデルの汎用性(画像での見え方に対するロバスト性)を向上させたり、識別モデルによる識別の精度を向上させたりすることも可能となる。

30

【0080】

(形態f)正解として1つの状態が対応付けられた1つの画像(学習用画像)に対し、少なくとも1つの見え方種別について、この1つの状態に対し予め設定された変更数だけの変更を行い、所定のサンプル数の学習画像群を生成する。

【0081】

この形態fによれば、例えば、(正解としての)ある1つの状態が対応付けられた学習用の画像サンプルを、必要とするサンプル数だけ準備することが可能となる。ここで、設定された変更数だけの変更は、例えば、見え方種別における見え方の「度合い」を段階的に変化させ、当該見え方種別について複数(変更数だけ)の互いに異なる度合いを有する変更画像を作成することによって達成される。

40

【0082】

ここで、1つの状態が対応付けられた画像のサンプル数が、他の状態が対応付けられた画像のサンプル数よりも小さい場合、当該1つの状態が対応付けられた画像について、当該他の状態が対応付けられた画像についてよりも大きい変更数だけの変更を行うことも好ましい。例えば、顔の表情がポジティブ、ニュートラル及びネガティブのうちのいずれの状態であるかを識別する表情識別モデルを構築する際、正解としてネガティブ状態である表情の顔画像サンプルが、他の状態に比べて比較的取得し難いとする。この場合、ネガティブ状態の元画像については、他の状態の元画像と比較して、例えば(図3(D)で説明した)顔の照度(明度)に係る見え方種別において、顔画像領域での明度(照度)の平均値をより細かく変化させ、より多くの変更画像を作成することができる。これにより、表情識別モデルに対し、より偏りの少ない良好な学習を行わせることも可能となる。

50

## 【0083】

(形態g) 識別モデルを構築する際、及びこの識別モデルを用いて状態を識別する際の両方で画像群を使用するケースにおいて、状態識別の際に使用される判定画像群を生成する場合、入力された1つの画像(識別対象画像)に対し、識別対象の見え方に係るいずれの種別についても、識別モデル構築の際に使用された学習画像群を生成する際に行われた種別についての変更と同じ方向への変更を行い、一方、学習画像群を生成する際に当該種別についての変更が行われなかった場合には変更を行わない。

## 【0084】

この形態gによれば、学習画像群と判定画像群とにおいて、いずれの見え方種別についても同方向への変更処理の施された画像が準備可能となる。すなわち、学習画像群と判定画像群との間で、見え方の変更の傾向を揃えることができる。その結果、例えば、所定の方向によって特徴付けられた(例えば、特に下向きの顔の表情を精度良く識別可能な)識別モデルを構築した上で、この識別モデルの当該特徴に合わせた画像入力を行うことによって、識別精度を向上させることも可能となるのである。

10

## 【0085】

図4は、本発明に係る変更画像作成・画像群生成処理の一実施形態を示すフローチャートである。なお、同図に示されたフローは、図2に示されたステップS104及びS204と、ステップS105及びS205との処理フローに対応している。

## 【0086】

(S301) 元画像に対する変更処理について、変更すべき見え方種別、並びに当該種別についての変更における方向及び度合いを設定する。

20

(S302) ここで、どの見え方種別について変更処理を行うかを決定し、設定された見え方種別について順次処理を進める。

## 【0087】

なお、設定された見え方種別、及び当該種別についての変更における方向及び度合い(属性)は、見え方種別一覧データベースに登録され、適宜参照されることも好ましい。ちなみに、本実施形態では、顔のはみ出しに係る種別、顔の向き・傾きに係る種別、装着物付加に係る種別、及び顔の照度(明度)に係る種別の4つの見え方種別が設定され、これらの見え方種別について順次処理が実施される。

30

## 【0088】

(S303a、S304a) 顔のはみ出しに係る見え方種別につき、入力画像(元画像)において抽出された顔(画像領域)が、画面からはみ出ているか否かを判定する。ここで、真の判定(はみ出ているとの判定)を行った場合、当該見え方種別については変更処理を行わない。一方、偽の判定を行った場合、顔のはみ出していない入力画像(元画像)に対し、所定の方向に所定量だけ顔(画像領域)を画面からはみ出させるセグメンテーション処理を実施する。

## 【0089】

(S303b、S304b、S304c) 顔の向き・傾きに係る見え方種別につき、入力画像(元画像)において抽出された顔(画像領域)が、上下左右のいずれかの方向に向いているか否か、及び顔が傾いているか否かを判定する。ここで、いずれについても偽の判定(正面向きであって傾いていないとの判定)を行った場合、所定の方向に所定の角度だけ顔を傾ける(例えば顔画像領域を回転させる)処理を行う。またはこの場合に、当該見え方種別について変更処理を行わないとしてもよい。一方、いずれかについて偽以外の判定を行った場合、顔の向き若しくは傾きを変化させる(例えば反転させる)処理を行ったり、所定の方向に所定の角度だけ顔を傾ける(例えば顔画像領域を回転させる)処理を行ったりする。

40

## 【0090】

(S303c、S304d) 装着物付加に係る種別につき、入力画像(元画像)において抽出された顔画像領域又はその付近に、所定の装着物(例えば眼鏡又は帽子)が存在するか否かを判定する。ここで、真の判定(所定装着物が存在するとの判定)を行った場合、

50

当該見え方種別については変更処理を行わない。一方、偽の判定を行った場合、所定の装着物の付されていない入力画像（元画像）に対し、顔画像領域又はその付近における所定の位置に所定の装着物を付加する処理を実施する。

【0091】

（S303d、S304e、S304f、S304g）顔の照度（明度）に係る種別につき、入力画像（元画像）において抽出された顔画像領域の全画素での明度（照度）の平均値が所定の基準明度（照度）範囲内であるか、当該範囲を超えて高すぎるのか、又は当該範囲を下回って低すぎるのかを判定する。ここで、当該範囲内であるとの判定を行った場合、当該見え方種別については変更処理を行わない。またはこの場合に、所定の明度（照度）変更処理を行ってもよい。一方、明度（照度）平均値が当該範囲を超えて高すぎる又は下回って低すぎるとの判定を行った場合、照度（明度）の高すぎる入力画像（元画像）に対し、明度（照度）の平均値が所定の基準明度（照度）範囲内となるように照度（明度）変更処理を実施する。

10

【0092】

（S305）設定した全ての見え方種別について、変更処理の判断を行ったか否か又は変更処理を行ったか否かの判定を行う。ここで、偽の判定を行った場合、変更処理（の判断）を行っていない見え方種別について当該処理を実施すべく、ステップS302に移行する。なお、1つの入力画像（元画像）に対し、複数の見え方種別についての変更処理を実施することも可能である。この場合、このステップS305において、すでに1つの見え方種別に関して変更処理の実施された入力画像について、残りの見え方種別に関する変更処理を受けさせる判断をすることになる。

20

【0093】

（S306）一方、ステップS305で真の判定を行った場合、作成された変更画像を含む学習画像群又は判定画像群を生成する。これで、変更画像作成・画像群生成処理が完了する。

【0094】

なお、当然に、本発明に係る変更画像作成・画像群生成処理は、以上に説明した実施形態に限定されるものではない。例えば、上述した形態a～形態gのうちの少なくとも1つを適用し、又は複数を組み合わせ、様々な処理形態を設定することが可能となる。いずれにしても、本発明に係る変更画像作成・画像群生成処理によれば、従来と比較してより少ない元画像をもって、十分な数の画像サンプルを準備することができる。または、従来と比較してより多量の画像サンプルを準備することも可能となるのである。

30

【0095】

図1に戻って、表情識別処理部114は、本実施形態において、識別モデル構築部114aと、表情判定部114bとを有する。このうち、識別モデル構築部114aは、生成された学習画像群を用いて学習を行い、（表情）識別モデルを構築する。この際、学習画像群に含まれる各画像サンプルには、（そもそも元画像の有していた）正解情報が付加されている。識別モデル構築部114aでは、この正解情報を参照し照合して学習を進めることになる。

【0096】

ここで、構築される識別モデルは、例えば、ディープラーニングの一種である畳み込みニューラルネットワーク（CNN, Convolutional Neural Network）を含む識別器とすることができる。なお当然に、本発明に係る学習画像群によって構築される識別モデルは、CNNを含む識別器に限定されず、状態識別の可能な他の機械学習手法に係る識別器としてもよい。さらには、例えば表情を判定する識別モデルとして、顔画像領域における目、鼻、口や、眉毛等の部位の特徴量を生成し、基準画像領域の特徴量との類似度に基づいて表情を判定する方式のものを採用することも可能である。

40

【0097】

一方、表情判定部114bは、生成された判定画像群を構築された識別モデルに入力し、判定画像群に含まれる変更画像、及び識別対象画像（元画像）における識別対象の状態

50

(表情)を判定し、判定状態(例えば表情がポジティブならば1、ニュートラルならば0、及びネガティブならば-1)並びに判定値(スコア)を出力する。なお、識別モデル構築部114a及び表情判定部114bのいずれか一方においては、生成された画像群を用いず、通常の学習用の多数の画像や識別対象画像を用いて処理を行うことも可能である。

【0098】

図5は、識別モデル構築部114aで構築される識別モデルの一実施形態を示す模式図である。

【0099】

図5に示すように、本実施形態において、識別モデル構築部114aで構築される識別モデルは、順伝播型の一つであるCNNに基づいて構成されている。このCNNは複数の畳み込み層を含んでいるが、この畳み込み層は、動物の視覚野の単純細胞の働きを模しており、画像に対しカーネル(重み付け行列フィルタ)をスライドさせて特徴マップを生成する畳み込み処理を実行する層である。この畳み込み処理によって、画像の解像度を段階的に落としながら、エッジや勾配等の基本的特徴を抽出し、局所的な相関パターンの情報を得ることができる。

【0100】

また、このCNNにおいて、各畳み込み層がプーリング層(サブサンプリング層)と対になっていて、畳み込み処理とプーリング処理とが繰り返し実施されることも好ましい。ここで、プーリング処理とは、動物の視覚野の複雑細胞の働きを模した処理であり、畳み込み層から出力される特徴マップ(一定領域内の畳み込みフィルタの反応)を最大値や平均値等でまとめ、調整パラメータを減らしつつ、局所的な平行移動不変性を確保する処理である。これにより、画像における多少のズレによる見え方の違いを吸収し、本来の特徴を捉えた適切な特徴量を獲得することができる。

【0101】

学習画像群生成部113bは、このCNNに対し学習を行わせるべく、生成した学習画像群を識別モデル構築部114aに出力するのである。具体的には、多数の元画像から生成された大量の学習画像群に含まれる画像サンプルを入力し、CNN内の複数の層のうち最終層を除いたいくつかの層分による多層ネットワークとしての反応を特徴量として出力し、この出力を正解と照合して、ニューロンの結合荷重やネットワーク構成のパラメータ等を生成・更新することにより学習を行う。

【0102】

通常、このようなCNNの学習は、例えば一般画像データベースに蓄積された大量の識別対象画像からなる大規模画像データセットを用いて行われる。これに対し、本実施形態では、元画像から学習画像群が生成された上でCNNへの入力が行われるので、通常必要とされる学習用画像のサンプル数に比べ、より少ない画像数(元画像数)で必要とされる大量の画像サンプルを準備することができるのである。

【0103】

図1に戻って、表情識別処理部114の表情判定部114bは、生成された判定画像群を構築された識別モデルに入力し、判定画像群に含まれる変更画像、及び識別対象画像(元画像)における識別対象の状態(表情)を判定する。ここで、構築された表情の識別モデルが、学習画像群に付加された正解情報として、例えば、ポジティブ、ニュートラル及びネガティブという表情に関する3つのカテゴリに対応する値(例えば1、0及び-1)で学習されている場合を説明する。この場合、表情判定部114bにおける各画像に対する判定状態(例えば表情がポジティブならば1、ニュートラルならば0、及びネガティブならば-1)は、この3つのカテゴリのうちで、対応する判定値(スコア)が最も高くなっているカテゴリの状態に決定される。なお、この判定値(スコア)は、対応するカテゴリのコンフィデンス(信頼度)を示す指標となっており、合計値が1となるように規格化されていることも好ましい。

【0104】

表情識別処理部114は、構築した識別モデルを用いて、生成された判定画像群に含ま

10

20

30

40

50

れる画像の各々についての状態（表情）を判定して、判定状態（例えば表情がポジティブならば1、ニュートラルならば0、及びネガティブならば-1）と、各カテゴリに対応する判定値とを出力し、これらの決定された判定値（スコア）に基づいて、最終的な状態識別を行う。例えば、これらの判定画像群の画像毎に決定された判定値（スコア）の平均値を、状態識別結果とすることも好ましい。

【0105】

また、変更態様として、

(a) 状態を判定された画像が変更前の元画像か否かに応じて、互いに異なる「重み」を予め設定し、この重みを用いて、判定画像群の画像毎に決定された判定値（スコア）の重み付け平均を行い、この重み付け平均結果を状態識別結果としてもよい。例えば、元画像の判定値に対する重みを、他の画像（変更画像）に対する重みよりも大きい値として重み平均処理を行うことによって、最終的な状態識別結果に、元画像の判定結果をより強く反映させることも可能となる。

10

【0106】

また、更なる変更態様として、

(b) 当該画像における変更された種別に応じて、互いに異なる「重み」を予め設定し、この重みを用いて、判定画像群の画像毎に決定された判定値（スコア）の重み付け平均を行い、この重み付け平均結果を状態識別結果としてもよい。例えば、顔の表情の判定結果が顔の向き・傾きの具合に最も強く影響を受けることが分かったとすると、顔の向き・傾きに係る見え方種別について変更した変更画像の判定値に対する重みを、他の見え方種別について変更した変更画像の判定値に対する重みよりも大きい値として重み平均処理を行ってもよい。これにより、顔の向き・傾き具合の強い影響を考慮した状態識別結果を得ることも可能となる。

20

【0107】

さらに、上記(a)及び(b)の両方に基づく「重み」を設定し、この重みによって状態識別結果を導出することも好ましい。また、各画像の判定値に対し、(重み)平均以外の演算処理、例えば多数決、を行って最終的な状態識別結果を決定することも可能である。

【0108】

ここで、以上に説明した状態識別結果（及び状態判定結果）は、直接例えばアプリケーション121に出力されて使用されてもよく、または、判定結果記憶部105に保存され、その後適宜、例えばアプリケーション121によって読み出されて使用されることも好ましい。アプリケーション121は、取得した状態識別結果（及び状態判定結果）を、例えば所定のアプリケーション・プログラムにおける表情判断データとして利用することができる。さらに、この利用結果（処理結果）や元の状態識別結果（及び状態判定結果）は、入出力制御部122を介して、タッチパネル・ディスプレイ106に表示されてもよく、通信インタフェース部101を通して外部の情報処理装置へ送信されてもよい。

30

【0109】

[状態識別システム]

図6は、本発明による状態識別システムの一実施形態を示す模式図である。

40

【0110】

図6によれば、本実施形態の状態識別システムは、サーバとしての表情識別準備装置3と、本発明による状態識別装置としてのスマートフォン5とを有している。

【0111】

このうち、表情識別準備装置3は、見え方度合い判定部312と、画像群生成部313と、識別モデル構築部314aを含む表情識別処理部314とを有する。これらの機能構成部は、図1に示したスマートフォン1における同名の機能構成部と同様の機能・構成を有している。したがって、表情識別準備装置3は、例えば外部の画像管理サーバ2から取得した学習用画像を元画像として学習画像群を生成し、生成した学習画像群を用いて識別モデルを構築し、構築した識別モデル（の情報）を、例えばスマートフォン5に送信する

50

ことができる。ちなみに、このような構成を有する本実施形態の表情識別準備装置 3 は、識別モデルに学習させるための学習画像群を生成するので、本発明による状態識別装置として把握することも可能である。

#### 【0112】

一方、スマートフォン 5 は、通信インタフェース部 501 と、カメラ 502 と、タッチパネル・ディスプレイ 503 と、画像管理部 511 と、見え方度合判定部 512 と、画像群生成部 513 と、表情判定部 514b を含む表情識別処理部 514 と、アプリケーション 521 と、入出力制御部 522 とを有する。これらの機能構成部は、図 1 に示したスマートフォン 1 における同名の機能構成部と同様の機能・構成を有している。したがって、スマートフォン 5 は、例えば、表情識別準備装置 3 から構築された識別モデル（の情報）を受信し、例えばカメラ 502 で撮影された識別対象画像を元画像として判定画像群を生成し、生成した判定画像群によって、取得した識別モデルを用いて状態の識別を行い、この状態識別結果をアプリケーション 521 で利用することができる。

10

#### 【0113】

以上説明したように、本実施形態では、識別モデルの構築（学習）は、表情識別準備装置 3 が実施している。したがって、スマートフォン 5 は、画像管理サーバ 2 等から大量の画像サンプルを取得しなくてもよく、また、処理負担の大きい識別モデル構築処理を実施する必要もない。その結果、スマートフォン 5 では、装置内で実行する情報処理量、及び装置内に取り込むべき情報量が格段に小さくて済む。言い換えれば、スマートフォン 5 は、携帯端末レベルのサイズ及び処理能力をもって好適な表情識別を実現可能とするのである。

20

#### 【0114】

なお、更なる他の実施形態として、スマートフォン 5 は、見え方度合判定部 512、画像群生成部 513、及び表情識別処理部 514 のいずれも備えておらず、表情識別準備装置 3 がこれらの機能構成部を全て備えていて、スマートフォン 5 は、表情識別準備装置 3 から状態識別結果（の情報）を取得して利用してもよい。このような実施形態では、表情識別準備装置 3 が本発明に係る状態識別装置として機能する。

#### 【0115】

ちなみに、上述したようなサーバ（表情識別準備装置 3）から出力された表情識別結果を享受する端末は当然、スマートフォンに限定されるものではない。例えば、タブレット型コンピュータ、ノート型コンピュータ、PC や、IOT（Internet Of Things）環境での使用に適したデバイスとしてのシンクライアント（Thin client）端末であってもよく、さらには、STB、サイネージ、ロボット等、種々の形態の端末を採用することが可能である。

30

#### 【0116】

以上、詳細に説明したように、本発明によれば、識別対象の見え方を変更し且つその状態を維持する処理で生成された画像を含む画像群を生成する。これにより、従来と比較して、種々の見え方をしているより多くの識別対象の画像を確実に準備することが可能となる。その結果、様々な状況下で種々の見え方をしている識別対象の画像であっても、当該状況や当該見え方による影響を調整し、より高い精度でこの識別対象の状態を識別することができるのである。

40

#### 【0117】

具体的に、本発明によれば、例えば、学習用の画像サンプル数が少ないため識別モデルの識別精度が低いままであるという従来の問題に対し、その学習用の画像サンプル数（さらには画像パターン数）を増やすことによって、高い識別精度を有する（さらには状況や見え方に対するロバスト性の高い）識別モデルを構築することが可能となる。また、識別モデルを用いた状態識別の際に、入力する画像数（さらには画像パターン数）を増やし、識別精度を向上させる（さらには状況や見え方に対するロバスト性を高める）ことも可能となるのである。

#### 【0118】

50

ちなみに、本発明に基づいて実現された高精度の表情識別結果を利用することによって、様々なサービスを提供可能なアプリケーション・プログラムを開発することもできる。そのようなアプリとして、例えば、この表情識別結果を利用して、対話している端末ユーザの感情（発話意図）を理解し、その応答内容を調整したり、当該ユーザとの対話内容をパーソナライズしたりすることが可能な対話AIアプリが挙げられる。

【0119】

以上に述べた本発明の種々の実施形態について、本発明の技術思想及び見地の範囲内の種々の変更、修正及び省略は、当業者によれば容易に行うことができる。以上に述べた説明はあくまで例示であって、何ら制約を意図するものではない。本発明は、特許請求の範囲及びその均等物によってのみ制約される。

10

【符号の説明】

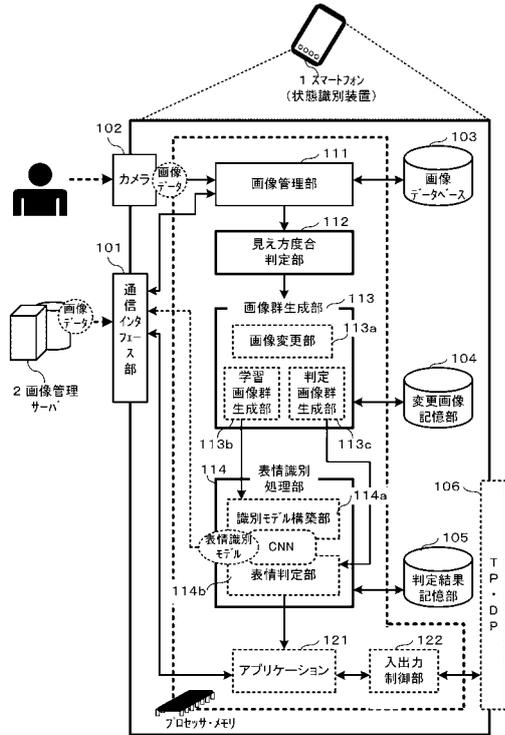
【0120】

- 1、5 スマートフォン（状態識別装置）
- 101、501 通信インタフェース部
- 102、502 カメラ
- 103 画像データベース
- 104 変更画像記憶部
- 105 判定結果記憶部
- 106、503 タッチパネル・ディスプレイ（TP・DP）
- 111、511 画像管理部
- 112、312、512 見え方度合判定部
- 113、313、513 画像群生成部
- 113a 画像変更部
- 113b 学習画像群生成部
- 113c 判定画像群生成部
- 114、314、514 表情識別処理部
- 114a、314a 識別モデル構築部
- 114b、514b 表情判定部
- 121、521 アプリケーション
- 122、522 入出力制御部
- 2 画像管理サーバ
- 3 表情識別準備装置

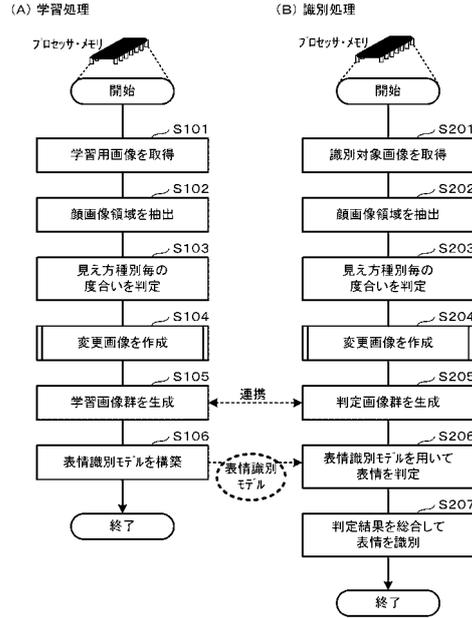
20

30

【図 1】



【図 2】



【図 3】

(A) 顔が画面からはみ出した画像



(B) 顔の向き・傾きの角度が大きい画像



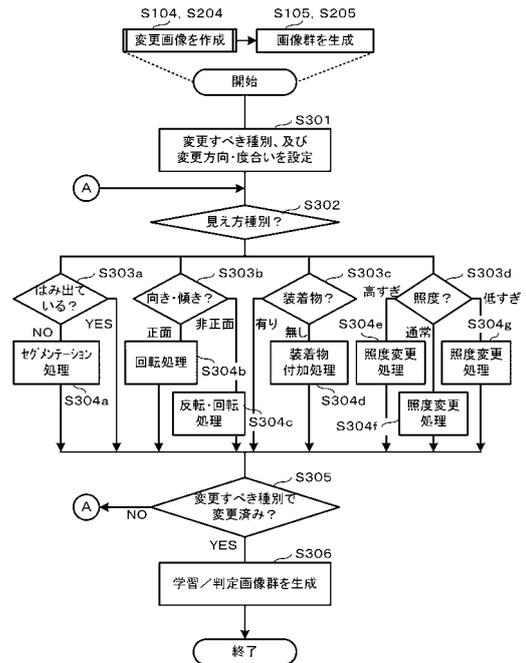
(C) 顔又は顔付近に装着物が付されている画像



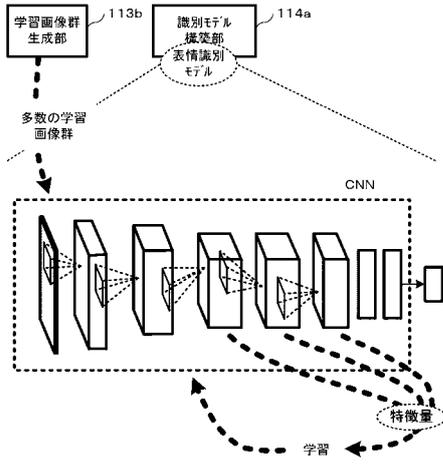
(D) 照度が高すぎる／低すぎる画像



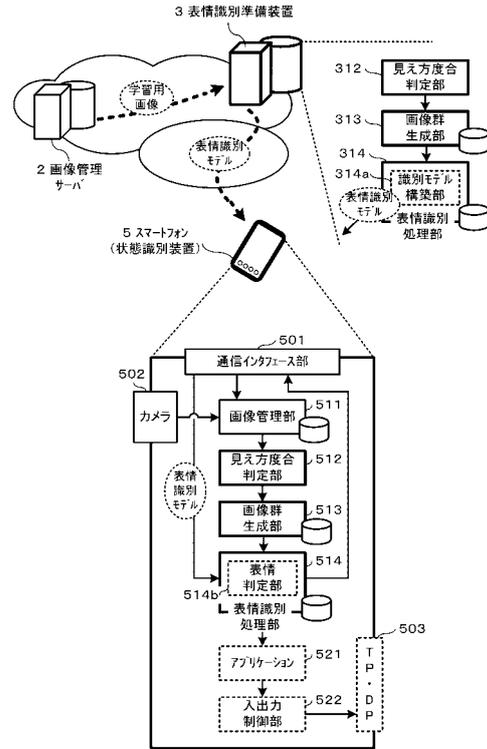
【図 4】



【 図 5 】



【 図 6 】



---

フロントページの続き

(72)発明者 矢崎 智基

埼玉県ふじみ野市大原二丁目1番15号 株式会社K D D I 総合研究所内

Fターム(参考) 5B057 BA02 CA08 CA12 CA16 CB08 CB12 CB16 CC01 CD03 CE09  
CE17

5L096 CA02 HA11 JA03 JA11 KA04 KA15