

【公報種別】特許法第17条の2の規定による補正の掲載
 【部門区分】第6部門第3区分
 【発行日】令和6年6月27日(2024.6.27)

【国際公開番号】WO2022/002961
 【公表番号】特表2023-531264(P2023-531264A)
 【公表日】令和5年7月21日(2023.7.21)
 【年通号数】公開公報(特許)2023-136
 【出願番号】特願2022-580296(P2022-580296)
 【国際特許分類】

10

G 0 6 T 7/00(2017.01)
 G 0 6 V 10/82(2022.01)
 G 0 6 V 40/16(2022.01)
 G 0 6 T 7/90(2017.01)
 G 0 6 T 7/60(2017.01)

【F I】

G 0 6 T 7/00 6 6 0 A
 G 0 6 T 7/00 3 5 0 C
 G 0 6 V 10/82
 G 0 6 V 40/16 Z
 G 0 6 T 7/90 A
 G 0 6 T 7/60 1 5 0 S

20

【手続補正書】
 【提出日】令和6年6月19日(2024.6.19)
 【手続補正1】
 【補正対象書類名】特許請求の範囲
 【補正対象項目名】全文
 【補正方法】変更

【補正の内容】
 【特許請求の範囲】

30

【請求項1】

顔の属性を決定する方法であって、
 ディープラーニング及び教師付き回帰を実行するネットワークモデルを使用して顔のソース画像を処理し、顔属性の各々について予測を出力することと、
 前記ネットワークモデルは、ディープラーニングを実行する残差ブロックを含む畳み込みニューラルネットワーク(CNN)モデルを含み、前記顔属性を予測するために、それぞれの分類器による分類のための共有特徴の特徴ベクトルを生成することを含むことを特徴とする方法。

【請求項2】

40

前記顔属性は、前記顔属性の固有の関係および異質性に従って複数のそれぞれの属性グループに関連付けられ、前記ソース画像は、前記属性グループのうちのそれぞれの1つのために構成されたそれぞれのネットワークモデルによって処理されることを特徴とする請求項1記載の方法。

【請求項3】

それぞれの前記属性グループのうちの1つは、色ベースの顔属性のための色属性グループであり、前記ソース画像は、色ベースのネットワークモデルによって処理されることを特徴とする請求項2記載の方法。

【請求項4】

それぞれの前記属性グループのうちの1つは、形状ベースの顔属性のための形状属性グル

50

ープであり、前記ソース画像は、形状ベースのネットワークモデルによって処理されることを特徴とする請求項 2 又は 3 に記載の方法。

【請求項 5】

前記顔属性のそれぞれについて、それぞれの分類器によって処理されるべき前記特徴ベクトルを複製することを含むことを特徴とする請求項 1 から 4 のいずれかに記載の方法。

【請求項 6】

前記ネットワークモデルが複数のそれぞれの分類器を備え、それぞれの前記分類器の各々が、1 又は複数の完全に接続された線形レイヤを備え、それぞれの前記分類器の各々が前記顔属性のそれぞれの 1 つの予測を出力として提供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方法。

【請求項 7】

複数のそれぞれの前記分類器は、前記顔属性を提供するために並行して実行することを特徴とする請求項 6 に記載の方法。

【請求項 8】

前記モデルは、予測レイヤのトリミングされた、適応され事前訓練された ResNet ベースの画像処理ネットワークモデルを含むことを特徴とする請求項 1 から 7 のいずれかに記載の方法。

【請求項 9】

画像検索、顔認識ならびに製品および/またはサービス推薦のうちのいずれか 1 つを実行するための前記顔属性を提供することを含むことを特徴とする請求項 1 から 8 のいずれかに記載の方法。

【請求項 10】

アプリケーションが前記顔属性に応答して前記製品および/またはサービス推奨を実行し、前記アプリケーションが、推奨される製品またはサービスに関連する効果をシミュレートするように前記ソース画像を修正して、ユーザエクスペリエンスに仮想的な試着を提供することを特徴とする請求項 9 に記載の方法。

【請求項 11】

顔画像の顔の属性を決定する方法であって、

残差ブロックを有する畳み込みニューラルネットワーク (CNN) ベースのバックボーンネットワークモデルを使用して顔のソース画像を処理して、共有特徴の特徴ベクトルを生成することと、

複数 (K 個) の分類器を使用して特徴ベクトルを処理し、K 個のそれぞれの顔属性が決定され、それぞれの分類器は、K 個のそれぞれの顔属性のうちの 1 つの予測を出力するために 1 又は複数の完全に接続された線形レイヤと予測ブロックとを備えることとを含むことを特徴とする方法。

【請求項 12】

前記顔の N 個のソース画像に対して前記処理ステップを実行するステップと、各顔属性の前記 N 個の予測から前記 K 個のそれぞれの顔属性の最終予測を決定するために予測解決ルールを使用することとを含むことを特徴とする請求項 11 に記載の方法。

【請求項 13】

第 1 のネットワークモデルが、色ベースの顔属性について前記ソース画像を処理して、色ベースの顔属性を予測するように構成された複数の前記分類器のそれぞれによって処理するための色ベースの特徴ベクトルを生成することを特徴とする請求項 11 又は 12 に記載の方法。

【請求項 14】

第 2 のネットワークモデルが、形状ベースの顔属性について前記ソース画像を処理して、形状ベースの顔属性を予測するように構成された複数の前記分類器のそれぞれによって処理するための形状ベースの特徴ベクトルを生成することを特徴とする請求項 11 から 13 のいずれかに記載の方法。

【請求項 15】

10

20

30

40

50

前記 K 個の分類器による並列処理のために前記特徴ベクトルを複製することを特徴とする請求項 11 から 14 のいずれかに記載の方法。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【発明の詳細な説明】

【相互参照】

【0001】

本出願は、2020年6月29日に出願され「改善された顔属性分類およびその使用のためのシステム及び方法」と題された米国仮出願第63/045,303号と、2020年10月5日に出願され「改善された顔属性分類およびその使用のためのシステム及び方法」と題されたフランス特許出願第2010128号と「改善された顔属性分類およびその使用のためのシステム及び方法」と題されたそれぞれの内容全体とが参照により本明細書に組み込まれる優先権を主張する。

【技術分野】

【0002】

本出願は機械学習を使用する画像処理に関し、より詳細には、顔属性分類のタスクに関する。顔画像が与えられると、顔属性分類は、所望の属性が顔のポートレートに存在するかどうかを推定することを目的とする。ダウンストリームアプリケーションは、様々な画像処理及び他のアプリケーションを含む。

【背景技術】

【0003】

顔属性分類 (Facial attribute classification、FAC) は、被写体の顔画像が与えられると、与えられた被写体の顔属性を分類しようとするタスクである。FACは、画像検索、顔認識、人物再識別および推薦システムを含む、多くのダウンストリームアプリケーションにとって非常に重要である。顔の画像を分析して、唇の大きさ及び形状、目の色等のような複数の属性および毛髪の色、性別等のような顔自体に厳密に限定されない他の属性 (一般に顔の属性として示される) を予測することができる。

【0004】

最近、畳み込みニューラルネットワーク (CNN) の優れた性能のために、ほとんどの最先端の FAC 手法は、顔属性を分類するために CNN を利用する。概して、これらの手法は、以下のように分類することができる: (1) 単一ラベル学習ベース (single-label learning based) の FAC 手法および (2) 複数ラベル学習ベース (multi-label learning based) の FAC 手法。単一ラベル学習ベースの FAC 手法は通常、顔画像の CNN 特徴を抽出し、次いで、サポートベクトルマシン (Support Vector Machine、SVM) 分類器によって顔属性を分類する。しかしながら、これらの手法は各属性を個々に予測し、従って、属性間の相関を無視する。対照的に、複数の属性を同時に予測することができる複数ラベル学習ベースの FAC 手法は、CNN の下位のレイヤから共有特徴を抽出し、CNN の上位のレイヤ上の属性固有分類器を学習する。

【発明の概要】

【0005】

実施形態によれば、顔属性予測およびその使用のための方法およびシステム及び/又はコンピュータデバイスを含む、ディープラーニング教師付き回帰ベースのモデルが説明される。使用の例は、画像から決定された顔属性予測に回答して修正された画像を提供するための拡張および/または仮想現実インターフェース (例えばコンピューティングデバイス又は方法を介して) である。一実施形態では、顔属性に一致する顔効果がインターフェースに適用されるように選択される。一実施形態では、顔効果がメイクアップ製品に関連するメイクアップ効果を含む。一実施形態では、電子商取引インターフェースが拡張およ

10

20

30

40

50

び/または仮想現実インターフェースを提供し、顔属性に応答するメイクアップ製品を推奨するための推奨コンポーネントを含む。

【0006】

また、実施形態によれば、コンピュータプログラム製品が提供され、その製品態様は命令を記憶する(例えば、非一時的である)記憶デバイスを備え、命令はコンピューティングデバイスのプロセッサによって実行されたときに、本明細書の任意のそれぞれの実施形態によるような方法を実行するようにコンピューティングデバイスを構成する。

【図面の簡単な説明】

【0007】

【図1】23の顔属性について複数の(顔)属性について属性値決定を行う際の人間の一致度の評価と、顔属性分類タスクにおける人間のパフォーマンスとを示すグラフである。

10

【図2】23の顔属性について複数の(顔)属性について属性値決定を行う際の人間の一致度の評価と、顔属性分類タスクにおける人間のパフォーマンスとを示すグラフである。

【図3】一実施形態によるディープラーニングネットワークアーキテクチャを示すブロック図である。

【図4】一実施形態によるコンピューティングシステム、例えばネットワークを示すブロック図である。

【図5】一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

【図6】A、Bは、一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

20

【図7】A、Bは、一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

【図8】一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

【図9】A、Bは、一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

【図10】A~Dは、一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

【図11】A~Cは、一実施形態によるユーザアプリケーションのグラフィカルユーザインターフェースのための表示画面またはその一部の図である。

30

【図12】本明細書の実施形態によるそれぞれの動作のフローチャートである。

【図13】本明細書の実施形態によるそれぞれの動作のフローチャートである。

【図14】本明細書の実施形態によるそれぞれの動作のフローチャートである。

【図15】本明細書の実施形態によるそれぞれの動作のフローチャートである。

【図16】本明細書の実施形態によるそれぞれの動作のフローチャートである。

【図17】本明細書の実施形態によるシステムのブロック図である。

【0008】

【0009】

【0010】

40

【0011】

【0012】

【0013】

本概念は添付の図面を参照して本明細書に記載される、その特定の実施形態を通して最も良く説明され、ここで、同様の参照番号は全体を通して同様の特徴を指す。本発明という用語は本明細書で使用される場合、以下に記載される実施形態の基礎をなす発明の概念を意味することを意図しており、単に実施形態自体を意味するものではないことを理解されたい。さらに、一般的な発明概念は以下に記載される例示的な実施形態に限定されず、以下の説明はそのような観点から読まれるべきであることを理解されたい。

【発明を実施するための形態】

50

【0014】

顔属性分類 (Facial attribute classification、FAC) は、被写体の顔画像が与えられると、与えられた被写体の顔属性を分類しようとするタスクである。FAC は画像検索、顔認識 (例えば2つの写真が同じ顔を示すかどうかを決定する)、顔位置特定および推奨システムを含む、多くのダウンストリームアプリケーションにとり重要である。

【0015】

本明細書の実施形態では、タスクが属性の所定のセットにわたる分類として定義される。各属性は各対象 (subject) を分類することができる潜在的なクラス (例えば属性値) のセットで記述される。例えば、肌のトーンの属性は、フェア (Fair)、ライト (Light)、ミディアム (Medium)、タン (Tan)、ディープ (Deep) 及び超ディープ (Very Deep) の合計6つのクラスを有する。従って、この特定の属性に関するタスクは、人物の画像が与えられ、6つのクラスのうちのどれが与えられた画像を最も良く表すかを予測する。

10

【0016】

属性の所定のセットは、色属性または形状属性の2つの異なるタイプに大まかに分類することができる。色属性の例としては、髪の色、肌のトーン及びアンダートーン (undertone) が挙げられる。形状属性の例としては、目の形状、目のしわ及び唇の大きさが挙げられる。

【0017】

最近、畳み込みニューラルネットワーク (CNN) の優れた性能のために、ほとんどの最先端の FAC 手法は、顔属性を分類するために CNN を利用する。概して、これらの手法は、以下のように分類することができる: (1) 単一ラベル学習ベース (single-label learning based) の FAC 手法および (2) 複数ラベル学習ベース (multi-label learning based) の FAC 手法。単一ラベル学習ベースの FAC 手法は通常、顔画像の CNN 特徴を抽出し、次いで、サポートベクトルマシン (Support Vector Machine、SVM) 分類器によって顔属性を分類する。しかしながら、これらの手法は各属性を個々に予測し、したがって、属性間の相関を無視する。対照的に、複数の属性を同時に予測することができる複数ラベル学習ベースの FAC 手法は CNN の下位レイヤから共有特徴 (shared features) を抽出し、CNN の上位レイヤ上の属性固有分類器を学習する。

20

【0018】

典型的な学術 (academic) の環境とは異なり、本明細書で主に説明されるシステム及びそのダウンストリームのタスクは、ファッション及び美容産業に主に関連する。一般的な学術データセット (general academic dataset) における属性は、現在のタスクに最小限の価値を提供する、より広い範囲の顔の特徴をカバーする傾向がある。対照的に、新しいテーラーメイドデータセット (tailor-made dataset) は、良く設計された属性タイプのためのソフトラベル (soft labels) を有する画像を記憶することで定義される。新しいデータセットは、それぞれが異なる環境下で撮影された1~5つの画像を有する3790の被験者を含む。全部で23個の属性があり、各属性は複数 (例えば6個) の個々のヒトアノテータ (human annotators) によってラベル付けされ、独立して作用し、更に記載されるようにソフトラベルを使用する。

30

40

【0019】

ヒトアノテータは、顔属性のインスタンスのための特定の値に常に合意するとは限らない。1つのアノテータは、値が平均であると結論付けても良く、別のアノテータは例えば値が高いと結論付けても良い。従って、顔属性分類などの分類タスクに関する人間のパフォーマンスを変更することができる。即ち人間は、異なる意見を有するか、又は特定の属性値について異なる結論に達する可能性がある。一実施形態では、属性データセットがグラウンドトゥルース (ground truth) 属性値に関するそれぞれの「意見」を集約するために、複数の人間からのアノテーション (annotations) を用いてコンパイルされる。

【0020】

50

このアプローチは、ハード又はシングルグラウンドトゥールズを使用する従来のデータセットアノテーションアプローチとは対照的である。例として、1つの属性は、軽度、中等度または重度のしわ値を有するしわに関する。この例は、値間の線形関係を示す。対照的なものは、顎の形状が丸い、尖っている又は角ばっている、である。これらは数学的に関連しているが、現在のタスクのための新しいデータセットを作成する際に実行されたことはアルゴリズムが各画像に対して使用する属性のための別個のグループ又はバケット (bucket) を定義することである。極端に多様な顎形状が与えられると、分類器アルゴリズムは各バケットに対して信頼性をもたらし、即ち最終製品の推奨は顧客の頬をどのように輪郭づけるかを決定するとき、人間が行う可能性が最も高いバケットを使用することができることを意味する。このタスクに対する従来のアプローチは画像に対する「正しい」出力が何であるかを理解するために、全てのアノテータが言ったことを平均化することによって、「ハードラベル」を使用する。代わりに、それは「ソフトラベル」アプローチを使用して、「正しい」答えがない可能性があるとして認識される。この手段はアノテータが不一致であっても、アルゴリズムはたとえそれが最も高い信頼性を有すると信じられたものに確定するとしても、全てのアウトプットを推奨する柔軟性を有する。

10

【0021】

図1, 2は、23の顔属性に対する顔属性分類タスクにおける人間の一致および人間のパフォーマンスを示すグラフである。

【0022】

人間の同意は、アノテータがどれだけ互いに同意するかを評価するために設定される。一実施形態では、6つのヒトアノテータがある。アノテーションを行う場合、全ての場合において、アノテータは互いに良く一致するわけではない。例えば髪の色アノテーションの場合、4人がブラウンと言うことができるが、2人が暗いブラウンとすることも、黒とすることもさへある。第1のグラフは一致を示し、範囲の中心は平均値である。範囲は、一致の標準偏差 (std dev) を示す。一致は多数決に基づいて計算されるので、4/2の上記の例では、一致がこの特定の属性 (例では髪の色) に関するこの特定の対象について66%である。3/3を50%、2/2/2を33%としてカウント。次いで、平均およびstd devは、データセット全体にわたって各属性について決定される。

20

【0023】

人間のパフォーマンスとは、人間自身がどれだけ良いかを評価することである。つまり、基本的には「どれだけ数が正しかったか」ということである。「正しい」の定義は、一実施形態について定義される要素である。実施形態では、それは多数決に従う。しかし、他の例は大多数が明確でない例 (例えば写真のためのアノテーションのセット) である。例えば髪の色の写真は、2クラスで3/3の投票を得たのか、3クラスで2/2/2の投票を得たのか 多数決はない。そのような場合、本実施形態では、「同じ投票を含む」および「同じ投票を含まない」と名付けられた2つのメトリックのセットが定義される。「同じ投票を含む」は多数決を有さない場合を含み、「同じ投票を含まない」は、明確な多数決を有さない場合を単に無視する。無視すると、これらの事例は、計算の分子または分母のいずれにも含まれない。それらは無視される。一方、「同じ投票を含む」は、全てのトップクラスを正しくカウントする。3/3の場合、どちらも正しいので、人間のパフォーマンスは100%である。2/2/2の場合、まだ100%である。2/2/1/1の場合、66%。

30

40

【0024】

表1に23の属性とその値を示す。

【表 1】

| | 属性 | 属性値 |
|----|--------------|--|
| 1 | 肌のトーン | フェア、ライト、ミディアム、タン、ディープ、非常にディープ |
| 2 | アンダートーン | ニュートラル、赤、黄 |
| 3 | そばかす | フレックル |
| 4 | 顔の形 | 楕円形、丸形、心臓、四角形、長形、ダイヤモンド |
| 4a | 顔の長さ | 長い、短い |
| 4b | 顎の形 | 丸形、角形、とがった |
| 5 | 眉の形 | アーチ、ソフトアーチ、ストレート、カーブ |
| 6 | 眉の太さ | 薄型、中型、厚型 |
| 7 | まつ毛の長さ | 短い、中ぐらい、長い |
| 8 | 鼻の種類 | 狭い、中ぐらい、広い |
| 9 | 目の色 | 青、グレー、グリーンヘーゼル、ブラウン、黒 |
| 10 | 目の形 | 丸形、アーモンド形 |
| 11 | 目のしわ | モノリッド、フード付き、ニュートラル/クリーズ付き |
| 12 | 目の角度 | 下降、中立、上向き |
| 13 | 目の幅 | ワイドセット、ニュートラル、クローズセット |
| 14 | 目の深さ | ディープセット、中立、突出 |
| 15 | 目のクマ | いいえ、はい |
| 16 | 髪の色 | なし、白、グレー、薄いブロンド、ブロンド、ブラウン、ダークブラウン、黒、苺色、銅、赤、赤褐色、その他 |
| 17 | 髪の質感 | ストレート、ウェーブ、カーリー、ちぢれ/アフロ、なし |
| 18 | 髪の長さ | 超短い、短い、中ぐらい、長い、超長い、なし |
| 19 | 額の大きさ | 短い、平均的、長い |
| 20 | 唇の大きさ | 薄型、中型、フル |
| 21 | 口の幅 | 狭い、平均的に広い |
| 22 | キューピッドの弓の可視性 | フラット、中ぐらい、ハイ/ハート |

表 1

10

20

30

40

【0025】

顔の形状は、顔属性 4 a (顔の長さ) 及び 4 b (顎の形) の予測からの信号を組み合わせることによって予測される二次属性である。例えば楕円形 = 長い (顔) + 丸形 (顎)、丸型 = 短い (顔) + 丸形 (顎) である。

【0026】

訓練目的のために、データは、それに応じて 2 2 7 4、7 5 8 及び 7 5 8 の被験者を含む訓練、検証および試験サブセットに分割される。

【0027】

見かけの顔アーキテクチャ予測問題は、ディープラーニングを用いたコンピュータ実装

50

のための教師あり回帰問題として投げかけられる。属性は、同じ個体からの 1 ~ 5 個の画像から評価される。

【 0 0 2 8 】

顔属性の固有の関係および異質性を効果的に活用するために、属性を異なるグループに分割することができる。一実施形態によれば、属性は、2つの別個のグループに分けられた：色属性（髪の色、目の色およびアンダートーン等）および形状属性（唇の大きさ、目のしわ及び鼻の種類等）。属性の各グループ上で類似のアーキテクチャを共有する2つのサブモデルが訓練された。推論中に、（それぞれの属性ごとの）予測は最後に集約される。便宜上、2つのサブモデル（1つは色用、1つは形状用）を集約して、顔属性分類のためのスーパーモデルを形成する。しかしながら、サブモデルのいずれか1つは、それぞれの属性のみが分類のために望まれる場合に有用である。

10

【 0 0 2 9 】

2つのサブモデル（例えば形状サブモデル及び色サブモデル）は同じ構成を共有するが、様々な前処理されたデータ（ここでは訓練の始まり以前の任意のデータ増強の手段の前処理、顔検出および顔切り取り、サイズ変更等を含むが、これらに限定されない）上で訓練される。カラーサイン（color signs）の場合、髪の色および肌のトーンのような属性はカラー抽出を導くために背景も必要とするので（例えば陰影環境において、明るい肌のトーンの画像は背景が切り取られた場合、暗い肌のトーンとして誤って分類される可能性がある）、データセットは形状属性の背景よりも多くの背景を有するように前処理される。形状モデルは最小限の背景で切り抜かれた顔について訓練されるが、カラーモデルはより多くの背景情報を含んでいた。

20

【 0 0 3 0 】

一実施形態では、2つのサブモデルにおける前処理は：

【 0 0 3 1 】

1. オリジナルの画像を R B G 空間に変換する；

【 0 0 3 2 】

2. 画像内に存在する顔を検出する。（任意の標準的な顔検出アルゴリズムが動作するはずである。）これは色に対してはオプションであるが、精度を向上させる；

【 0 0 3 3 】

3. フェイスボックスのサイズ変更 - カラーモデル：検出されたフェイスボックスを 1 . 7 倍に拡大する、シェイプモデル：検出されたフェイスボックスを 1 . 2 倍の幅と高さ

30

【 0 0 3 4 】

4. 手順 3 . の画像をモデルの入力サイズにサイズ変更する。

【 0 0 3 5 】

従って、データセットは、カラー関連サイン（color related signs）及び形状関連サイン（shape related signs）をバケット化することにより、2つのアルゴリズムを訓練するために使用される。これは、関連するサインから学習されたアルゴリズムを意味する。例えば明るい毛髪を有する人々は青い目を有することが多く、暗い皮膚を有する人々は褐色の目を有することが多い。しかし、青い目を持つ人は、顔の形にほとんど相関関係がないかもしれない。各アルゴリズムの結果として得られるコードは、関連するサインの関連する計算を再利用し、関連するサイン間を接続することによって、非常に小さく、非常に高速である。2つのアルゴリズムは一緒になって、ユーザの顔特性の完全なピクチャを作り出すが、互いに過度に複雑にならない。

40

【 0 0 3 6 】

図 3 は、ディープレニングネットワークアーキテクチャを示すブロック図である。3 0 0 ネットワークは、入力レイヤ 3 0 4 においてソース画像を処理するための畳み込みニューラルネットワーク（CNN）3 0 2 を備える。一実施形態では、CNN 3 0 2 が複数の残差ブロック（residual blocks）3 0 6 を有する残差ネットワークベースのバックボーンを使用して、共有特徴を抽出するように構成される。例として、残余ネットワー

50

クベースのバックボーンは、ResNet ("Deep Residual Learning for Image Recognition", He, Kaiming et al, 2015-12-10, URL: arxiv.org/abs/1512.03385) を使用して構成される。一実施形態では、例えばImageNet(Stanford Vision Lab, Stanford University(URL image-net.org)から入手可能な組織化された画像のデータベース)上で事前訓練されたモデルが、バックボーンネットとして適合され、使用される。そのようなモデルの最後のいくつかの予測レイヤは手元の目的のために、例えば必要に応じて共有特徴データを生成するためのエンコーダとして、モデルを適応させるために除去される。

【0037】

(例えば平均プーリング(average pooling)を使用して)平坦化された特徴ベクトル308が、バックボーンネット302から取得される。例えば一実施形態では、ResNet18ベースのバックボーン特徴ネットから、出力はサイズ224×224の画像に対する次元7×7×512のベクトルを有する。平均プーリングは、一実施形態では512個の位置の各々における7×7個の項の平均をとることによって、1×512の平坦化ベクトルを提供する。

【0038】

特徴ベクトル308は、K個の顔属性の各々に対する複数(K個)の分類器310による処理のために(例えば並行して)複製される。各個々の分類器(例えば312₁、312₂、...312_K)は、1又は複数の完全に接続された線形レイヤ(314₁、314₂、...314_K)及び予測ブロック(316₁、316₂、...316_K)を備え、各属性の最終予測を出力する。形状モデルでは、各分類器ヘッド(classifier head)が整流線形ユニット(rectified linear unit, ReLU)活性化を有する2つの完全に接続された線形レイヤからなる。隠れレイヤ(hidden layer)サイズは50である(この手段、特徴サイズがNであると仮定すると、第1のFCレイヤは特徴を50のサイズに絞り込み、次いで、第2のFCレイヤはサイズ50の特徴に作用し、目標出力サイズに等しいサイズを有する特徴ベクトルを出力する)。カラーモデルでは、各分類器ヘッドが1つのFC線形レイヤのみからなる。(サイズNの特徴をターゲット出力サイズに直接変換する)。典型的には、モデルがより深い(即ちより多くのレイヤが存在する)ほど、モデルの容量はより大きくなる。形状属性は、完全に接続されたレイヤの数の差を説明する学習困難性の点で、色属性よりも比較的複雑である。

【0039】

訓練時および/または推論時の使用中に、一実施形態では、モデルが一度にN個の画像のグループを処理するために使用される。出力は(N, K)の次元を有し、Nは入力画像の数であり、Kは属性の数である。出力行列の各要素は整数であり、O_{i, j}は、i番目の画像およびj番目の属性について、システムが画像を予測するクラスが属することを表す。一実施形態では、Nは学習レートのための訓練ハイパーパラメータ(training hyper-parameter)である。N個の画像の各々は、同じ又は異なる顔を含む。このモデルは、同一性に関するコンテキストを有しておらず、同じ顔の複数の画像にわたって矛盾する予測結果を解決するものでもない。

【0040】

それぞれの属性についての分類器の各々の中で、最終出力レイヤの前のレイヤは、可能な属性値のセットにわたって(スコア値のよう)予測のセットを生成する。一実施形態では、各予測が属性値の「真」に関する相対的尤度である。3つの属性値のセットについて、予測のセットは20、20、60であり得るが、多くの場合、予測は異なるスケールにわたって正規化される。最終的な予測は予測値のセット(例えば60に関連付けられた属性)の最大値から決定される。

【0041】

上述のように、教師あり学習に使用される注釈付きデータセットは、6人の人間のアノテーションが独立して行動するアノテーションを有する。上述のように、特定の属性に対する注釈は、各注釈によって同一ではない。「真」解決ルールは、アノテーション値(ソフト

10

20

30

40

50

ラベル)の中から選択するために使用される。1つのルールは最も一般的な予測ルールであり、一連の予測から最も一般的に予測される属性を選択する。1つのルールは、上位2つの最も一般的な予測ルールである。

【0042】

カラーモデルの場合、一実施形態では、「真」が「最も一般的な投票」として扱われる。これは、目の色に対する以下のアノテーションを有することを意味する：「青」、「青」、「青」、「灰色」、「灰色」、「茶色」：モデルが「青」を予測する場合、最も一般的な投票との整合が存在する。モデルが「青」または「灰色」を予測する場合、上位2つの最も一般的な投票との整合が存在する。アノテーションが「青」、「青」、「灰色」、「灰色」、「茶色」、「黒」である場合、「青」または「灰色」のいずれかが最も一般的な投票である（および上位2つの最も一般的な投票）。

10

【0043】

表2は、訓練された色属性を予測するためのモデルの出力および評価を示す。「*」でアノテーション付けされた属性出力は、モデルによって予測されることがほとんどない出力を表す。一実施形態では、追加の訓練データを利用して、そのような予測を改善する。一実施形態では別のアプローチがこれらの粒度予測を、より一般的な（および関連する）予測と一緒に、又はバンドルすること、予測の粒度または細かさを低減することを含む。例えば一実施形態では「ストロベリー」、「レッド」及び「オーバーン」は共バンドルされる。

20

【表2】

| | 最も一般的な投票に一致するパーセンテージ | 上位2つの最も一般的な投票に一致するパーセンテージ | 出力 |
|---------|----------------------|---------------------------|--|
| 肌のトーン | 0.84180 | 0.95508 | フェア、ライト、ミディアム、タン*、ディープ、非常にディープ* |
| アンダートーン | 0.64355 | 0.89160 | ニュートラル、赤、黄 |
| 髪の色 | 0.70752 | 0.87451 | 銅*、ダークブラウン、苺色*、黒、ブラウン、未定、ブロンド、薄いブロンド、赤褐色*、白、赤*、その他*、グレー、なし |
| 目の色 | 0.69531 | 0.88281 | 黒、ブラウン、緑、グレー、ヘーゼル*、青 |

30

40

表2

【0044】

50

形状モデルでは、最も一般的な投票もグラウンドトゥールースとして扱われた。2つの評価基準（予測解決ルール（prediction resolving rule））が使用された：「多数決に一致するパーセンテージのみ」及び「最も一般的な投票に一致するパーセンテージ」。

【0045】

「最も一般的な投票に合致するパーセンテージ」は全ての最大投票クラスが「グラウンドトゥールース」として扱われるときに、試験セット内の全てのサンプルに対する精度を計算する。

【0046】

「多数決のみに一致するパーセンテージ」は、1つの多数決クラスのみを有するサンプルについての精度を計算するだけである。

【0047】

表3は、訓練された形状属性を予測するためのモデルの出力および評価を示す。

【表3】

| | パーセンテージ一致多数決のみ（即ち同じ投票を含まない） | 最も一般的な投票に一致するパーセンテージ | 出力 |
|----------------|-----------------------------|----------------------|---------------------------|
| 眉の形 | 0.6717 | 0.7124 | ストレート、カーブ、アーチ（+ソフト） |
| 目の形 | 0.9333 | 0.9406 | 丸形、アーモンド形 |
| 目のしわ | 0.7543 | 0.8021 | モノリッド、ニュートラル/クリース付き、フード付き |
| 顔の長さ (顔の形1) | 0.8459 | 0.8668 | 長い、大きい |
| 顎の形 (顔の形2) | 0.6604 | 0.6953 | とがった、丸形、角形 |
| 唇の大きさ | 0.8346 | 0.8377 | 薄型、中型、フル |

表3

【0048】

図4は、一実施形態において、顔属性分類に応答する推奨アプリケーションを提供するコンピューティングシステム400を示す。アプリケーションは、電子商取引アプリケーションである。更にディープラーニングネットワークモデル（deep learning network model）を使用すること等による画像処理技法によって、アプリケーションは、顔効果シミュレーションを提供して、推奨をコンテキスト的に提示する。アプリケーションは、仮想および/または拡張現実体験をユーザに提供する。

【 0 0 4 9 】

システム 4 0 0 には、ユーザコンピューティングデバイス 4 0 2 が示される。ユーザコンピューティングデバイス 4 0 2 は、スマートフォンの形態で示される。タブレット端末、パーソナルコンピュータ等の他のフォームファクタが有用である。ユーザコンピューティングデバイス 4 0 2 は、アプリケーション 4 0 4 等の電子商取引アプリケーションに有用な、Apple Inc (登録商標)、Samsung Electronics Co, Ltd. (登録商標) 等から入手可能なような消費者向けデバイスである。一実施形態では、デバイス 4 0 2 が販売員などによって操作される製品カウンターキオスクデバイスである。

【 0 0 5 0 】

一実施形態では、ユーザコンピューティングデバイス 4 0 2 がプロセッサ (CPU、GPU など)、記憶デバイス、1 又は複数の通信サブシステム又はデバイス、ディスプレイ、入力デバイス、出力デバイス等のうちの少なくともいくつかを含む、通信のために結合された様々な構成要素を有するコンピューティングデバイスを備える。一実施形態では、ディスプレイがタッチ又はジェスチャー可能なもの (gesture enabled) であり、入力デバイスはマウス、マイクロフォン、カメラ、キーボード、ボタン等のポインティングデバイスを含み、通信デバイスは有線または無線通信のいずれかを提供し、短距離および/または長距離の通信が可能である。一実施形態では、通信デバイスが衛星ベースのロケーションサービスを提供する等のためにロケーションデバイスに結合する。一実施形態では出力デバイスがスピーカ、ライト、振動/触覚デバイスなどを含み、様々な構成要素は 1 又は複数の通信バスまたは他の構造を介して結合される。

【 0 0 5 1 】

一実施形態では、記憶デバイスは、例えば実行されたときにコンピューティングデバイスの動作を構成する命令およびデータを記憶する。命令は、オペレーティングシステム、アプリケーション等を定義する。データは、顔の 1 つ以上のソース画像を含む。アプリケーション 4 0 4 に加えて、一実施形態では、電子メール、テキストメッセージ、インスタントメッセージ、ショートメッセージサービス (SMS) メッセージ等のメッセージ (データ) を通信するための 1 又は複数のアプリケーションがある。一実施形態では、通信が画像またはビデオなどの添付ファイルを含む。

【 0 0 5 2 】

一実施形態では、アプリケーション 4 0 4 がユーザコンピューティングデバイス 4 0 2 の一部として、又は、それに結合された、ディスプレイデバイス (例えば 4 0 8) による表示のための 1 又は複数の画面 (例えば 4 0 6) を含むグラフィカルユーザインターフェースを提供する。一実施形態では、ユーザコンピューティングデバイス 4 0 2 及びアプリケーション 4 0 4 がジェスチャ及び/又は音声であるか又はその他が有効である。一実施形態では、アプリケーション 4 0 4 がネイティブアプリケーションであり、ユーザデバイス及びそのオペレーティングシステムのために特に構成され、ダウンロードされ、その上に記憶されるか、又は、ブラウザベースのアプリケーションの形態など、より多くのデバイスに不可知 (agnostic) であり、及び/又は、ウェブサービスとして提供される。

【 0 0 5 3 】

一実施形態では、アプリケーション 4 0 4 が例えばインターネットプロトコル (IP) を使用して、ネットワーク 4 1 0 を介して、電子商取引サーバ 4 1 2 及び画像処理サーバ 4 1 4 等の 1 又は複数のサーバと通信する (デバイス 4 0 2 を使用して)。システム 4 0 0 は簡略化されており、支払い及び他のサービスは示されておらず、1 つのユーザコンピューティングデバイスのみが示されている。

【 0 0 5 4 】

一実施形態では、サーバ 4 1 2 及び 4 1 4 が (1 つのそのようなデバイスが使用され得るが) 既知の適切なリソースを有するそれぞれのコンピューティングデバイスによって提供される。適切なコンピューティングデバイスは、プロセッサ (CPU、GPU 等)、記憶デバイス、1 又は複数の通信サブシステム又はデバイス、ディスプレイ、入力デバイス、出力デバイス等のうちの少なくともいくつかを含む、通信のために結合された様々な構

成要素を有する。一実施形態では、ディスプレイがタッチ又はジェスチャ可能なものであり、入力デバイスはマウス、マイクロフォン、カメラ、キーボード、ボタンなどのポインティングデバイスを含み、通信デバイスは有線または無線通信のいずれかを提供し、短距離または長距離の通信が可能である。一実施形態では通信デバイスが衛星ベースのロケーションサービスを提供するなどのロケーションデバイスに結合し、出力デバイスはスピーカ、ライト、振動/触覚デバイス等のいずれかを含む。様々な構成要素は、1又は複数の通信バス又は他の構造を介して結合される。一実施形態では、記憶デバイスは、例えば実行されたときにコンピューティングデバイスの動作を構成する命令およびデータを記憶する。

【0055】

10

本実施形態では、アプリケーション404がメイクアップ製品のための電子商取引アプリケーションである。アプリケーション404は顔属性分析を取得するためにソース画像を提供し、ソース画像から決定された顔属性に応答して製品推薦を受信するように構成される。サーバ414はソース画像を(サーバ412への通信なしにデバイス402から、又はサーバ412を介して)受信し、顔の特徴を決定するための方法を実行するように構成される。

【0056】

サーバ414はディープラーニング及び教師付き回帰(supervised regression)を実行するネットワークモデルを使用して顔のソース画像を処理し、顔属性の各々について予測を出力する。

20

【0057】

一実施形態では、ネットワークモデルがディープラーニングを実行して、顔属性を予測するためにそれぞれの分類器による分類のための共有特徴の特徴ベクトルを生成する残差ブロックを有する畳み込みニューラルネットワーク(CNN)モデルを備える。特徴ベクトルの複製は、顔属性のそれぞれの1つについてそれぞれの分類器によって処理されるようにされる。一実施形態ではネットワークモデルが複数のそれぞれの分類器を備え、それぞれの分類器の各々は、1又は複数の完全に接続された線形レイヤを備え、それぞれの分類器の各々は顔属性のそれぞれ1つの予測を出力として提供する。複数のそれぞれの分類器は、一実施形態では顔属性を提供するために並行して実行される。

【0058】

30

一実施形態では、顔属性が顔属性の固有の関係および異質性(heterogeneity)(例えば色または形状など)に従って、複数(例えば2つ)のそれぞれの属性グループに関連付けられる。一実施形態では、サーバ414が、そのような手法がソース画像を処理するために、属性グループのそれぞれの1つのために構成されたそれぞれのネットワークモデルを使用するように構成される。

【0059】

一実施形態では、それぞれの属性グループのうち1つは、色ベースの顔属性のための色属性グループであり、ソース画像は色ベースのネットワークモデルによって処理される。一実施形態では、それぞれの属性グループのうち1つは形状ベースの顔属性のための形状属性グループであり、ソース画像は形状ベースのネットワークモデルによって処理される。

40

【0060】

一実施形態では、ネットワークモデルが予測レイヤのトリミングされた、適応され事前訓練されたResNetベースの画像処理ネットワークモデルを含む。

【0061】

サーバ414は、本実施形態では製品推薦を提供するために使用される顔属性を提供するが、他の実施形態では、サーバが画像検索、顔認識、個人再識別ならびに製品および/またはサービス推薦などのうちのいずれか1つを実行するアプリケーションによって使用するための顔属性を提供する。

【0062】

50

一実施形態では、サーバ414が残差ブロックを有する畳み込みニューラルネットワーク(CNN)ベースのバックボーンネットワークモデルを使用して顔のソース画像を処理し、共有特徴の特徴ベクトルを生成し、複数(K個)の分類器を使用して特徴ベクトルを処理し、K個のそれぞれの顔属性が決定され、それぞれの分類器は、K個のそれぞれの顔属性のうち1つの予測を出力するために1又は複数の完全に接続された線形レイヤと予測ブロックとを備える。一実施形態では、第1のネットワークモデル(例えばサブモデル)は色ベースの顔属性についてソース画像を処理して、色ベースの顔属性を予測するように構成された複数の分類器のそれぞれによって処理するための色ベースの特徴ベクトルを生成し、第2のネットワークモデルは形状ベースの顔属性についてソース画像を処理し、形状ベースの顔属性を予測するように構成された複数の分類器のそれぞれによって処理するための形状ベースの特徴ベクトルを生成する。

10

【0063】

ユーザコンピューティングデバイスがソース画像を(例えばカメラまたは他のアップロードから(例えば記憶デバイスから))提供するクライアント/サーバモデルとして示されているが、サーバ414は顔属性検出サービスを実行し、一実施形態ではユーザコンピューティングデバイスが十分にロバストである場合、分類を実行するためにディープラーニングモデルを用いて構成される。

【0064】

一実施形態では、サーバ412が電子商取引サービスのための製品推奨を提供する。サーバ412は、製品推奨の少なくともいくつかを購入するための電子商取引ショッピングサービスを提供する。

20

【0065】

一実施形態では、サーバ412が顔のソース画像から決定された複数の顔属性を受信することによって、複数の顔属性を生成するために顔属性分類ネットワークモデル(facial attribute classifying network model)を使用して処理されたソース画像を受信することと、少なくともいくつかの顔属性を使用して、製品に適した顔属性と関連付けて製品を記憶するデータストアから少なくとも1つの製品を選択することと、製品を購入するために電子商取引インターフェースにおいて提示するための推奨として少なくとも1つの製品を提供することと、などによって、電子商取引サービスのために製品を推奨するように構成される。

30

【0066】

一実施形態では、顔属性がサーバ414及び/又はサーバ414について説明した方法を使用して決定される。

【0067】

一実施形態では、製品がメイクアップ製品を含む。メイクアップ製品は所定のメイクアップルック(make-up looks)を定義するために、データストア(例えばサーバ412に結合されたデータベース416)内で様々に関連付けられる。ルックは、特定の効果を有するメイクアップから定義され、及び/又は特定の技術を使用して適用される。各ルックは異なるタイプのメイクアップから(例えば含むように)定義される。タイプは顔(頬、鼻、顎、額の線など)の皮膚の領域(領域)または眉毛およびまつ毛などの毛を含む目、唇および眉毛などの領域に関連する。顔の領域(皮膚)、目の領域、眉の領域、唇の領域などは、説明されるように、形状、色を含む顔属性に関連付けられる。それぞれの領域のためのそれぞれのメイクアップ製品は、例えば属性のための適合性尺度(suitability measure)を使用することによって、それぞれの属性に関連付けられる。測定値は、2値、例えば「良い」又は「良くない」であっても良く、又は、適合性測定値はより粒状であっても良い(例えば1~10のスケール又は他のスケールでランク付けされる)。一実施形態では、外観が顔の複数の領域(例えば全ての領域)に対するそれぞれの製品を含む。ルックのための製品は顔の属性に関連する尺度を有する(例えば唇の形状および色の属性に対する唇の製品、眉の属性に対する眉の製品など)推奨機能はユーザ画像の画像分析から決定された顔の属性を、それぞれのルックにグループ化された製品に一致させる

40

50

ために、ルール又は他の方法を使用することができる。一実施形態では、顔属性へのマッチングが潜在的な外観を識別する。2つ以上のルックが、初期マッチングから識別され得る。識別されたルックは例えばそれぞれのルックについての顔属性マッチのカウント、及び/又は、マッチのスケールを評価する顔属性マッチの品質に基づいてランク付けされ得る。ランキングは、ルックの提示を順序付けるため、及び/又は、スコアを提供するために使用され得る。一実施形態では、外観を選択し、顔属性情報を使用して、外観を達成するために使用可能なマッチング製品を決定することができる（例えば顔に組み合わせたとき）。ランク付けは、外観を達成するために使用されるそれぞれの製品を順序付けまたはスコア付けするために使用され得る。

【0068】

10

一実施形態では、サーバ414が所定のメイクアップルックのうちの1つの識別を受信する。顔属性のうちの少なくともいくつかを使用するとき、サーバ412は少なくとも1つの製品を選択するときに、所定のメイクアップルックのうちの1つに応答する。

【0069】

一実施形態ではメイクアップ製品の各々が複数のメイクアップタイプのうちの1つに関連付けられるとき、サーバ412はメイクアップタイプの各々について、顔属性に応答して、推奨を定義するために少なくとも1つの製品を選択する。上述のように、メイクアップタイプは、顔製品タイプ、目製品タイプ、眉製品タイプ、及び唇製品タイプを含む。

【0070】

一実施形態では、サーバ412が推奨製品を使用する技術を推奨する。

20

【0071】

一実施形態では、サーバ412がアプリケーション404のために（ユーザ）コンピューティングデバイス402にその出力を提供する。出力は、電子商取引サービスのグラフィカルユーザインターフェースなどを介してディスプレイ408を介して提示される。

【0072】

一実施形態ではユーザのルック又は個々のメイクアップ製品の知覚を向上させるために、ソース画像は顔効果を提示するために画像処理等によって修正される。ソース画像は、サーバ412によって、又はサーバ414によって、又は本明細書の実施形態で説明されるようなユーザコンピューティングデバイス402によって修正される。ディープラーニングネットワークモデルは、ソース画像を修正するか、又は他の画像処理技術が本明細書の実施形態で説明されるように使用される。ソース特徴を維持しながら、ある領域（domain）から別の領域に変換するように、ソース画像を修正するための1つのそのようなディープラーニングネットワークモデルは、敵対的生成ネットワーク（GAN）ベースのモデルである。

30

【0073】

一実施形態では、アプリケーション404がユーザコンピューティングデバイス412が現実をシミュレートするための方法を実行することを可能にする。方法は、顔のソース画像から複数の顔属性を決定することと、顔属性分類ネットワークモデルを使用してソース画像を処理することと、顔属性のうちの少なくとも1つに応答して、顔への適用のための少なくとも1つの顔効果を決定することと、表示のために少なくとも1つの顔効果を顔に適用するためにソース画像を処理することとを含む。一実施形態では、アプリケーション404がサーバ412及び414の一方または両方と通信して、そのために実行される方法ステップを有する。ユーザコンピューティングデバイス402は一実施形態ではカメラ418を使用してソース画像（例えば420）を生成する。

40

【0074】

一実施形態では、少なくとも1つの顔面効果が顔面に適用される少なくとも1つのメイクアップ製品および/または技法のシミュレーションである。一実施形態では、他の顔面効果がプロテーゼ、形成手術またはアイウェア等の適用を含む。

【0075】

様々な実施形態では、少なくとも1つの顔効果を適用するためにソース画像を処理する

50

ことは顔効果を適用するように構成されたディープラーニングネットワーク（例えばGANベース）を使用すること、又は、他のディープラーニングモデル若しくは他の画像処理技法が使用されることを含む。限定ではなく例として、（クライアント側）コンピューティングデバイスにおいて構成された拡張現実（AR）仮想的な試着のための方法が、ソース画像上の推奨製品またはサービスに関連する効果をシミュレートするために使用され、仮想的な試着のユーザエクスペリエンスを与える。要約すると、メイクアップ製品を表すレンダリング値（例えば赤、緑、青（RGB）色および色不透明度）と同様に、カメラフレーム又は写真（例えばソース画像として）が、方法への入力として受け取られる。ソース画像は、顔追跡を使用して、ソース画像内のユーザの顔の周りのキーランドマーク（key landmarks）を決定するために処理される。これらのランドマークを使用して、メイクアップ製品に対する顔の領域が、唇関連製品のためのユーザの唇のようにマスクアウト（masked out）される。次に、レンダリング値を使用して、マスクアウトされたエリア内のユーザの顔にメイクアップを描画する。一実施形態では、実際の製品がどのように見えるかを現実的にレンダリングするために、ソース画像から照明をサンプリングすること、ブレンディング（blending）することなどの他のステップが含まれる。最後に、メイクアップ又は顔効果を含む変更されたソース画像の結果が、表示可能な出力として返される。

10

【0076】

本実施形態において、図5、6A、6B、7A、7B、8、9A、9B、10A、10B、10C、10D、11A、11B及び11Cは、アプリケーション404のグラフィカルユーザインターフェースのための画面またはその一部を示す。

20

【0077】

本実施形態では、図5がアプリケーションの開始画面500を示し、このコンテンツはユーザコンピューティングデバイス402のディスプレイ画面408上に全てが適合する（全てが表示可能ではない）わけではない。部分は、アプリケーションの概要を提供する部分502等、知られているようにスクロール可能である。制御504はアプリケーションを進める（例えば開始する）ために提供される。

【0078】

本実施形態では、図6Aがソース画像（例えば420）を取得するための初期画面600（例えば「写真撮影指示画面」）を示す。画像は、カメラ418を介した「自撮り」である。「撮影」制御が進み、撮影インターフェースを開始する（図7参照）。アップロード制御604は、デバイス402のローカル記憶デバイスから、又は別様にスクリーン610（例えば「画像取得のためのアップロード/オプション」）及び図6Bを介して画像を取得するように、アップロードインターフェースに進む。

30

【0079】

本実施形態において、図7A及び図7Bは顔704の画像をキャプチャする（例えばカメラから受信する）ための撮影インターフェース（例えばそれぞれスクリーン700及び710）を示す。画像キャプチャ制御702は、カメラ418を呼び出して、ソース画像420を取得する。背面（反対向き）カメラモードに切り替えるためのものなど、他の制御は図示されていない。実施形態では、ユーザが顔を適切に位置決めするのを支援するため、メイクアップの摩耗を最小限にするため、髪または他の物体が顔を閉塞するのを最小限にするため及び適切な照明を使用するため等の命令が提示される。命令の例は「顔を位置付ける」706であり、これは、画像が撮影される前に顔のプレゼンテーションをオーバーレイすることができる。画像がキャプチャされると、顔の位置および照明を確認し、リテーク（図示せず）を要求するための初期処理が実行される。本実施形態では、ソース画像420が説明したように顔属性を決定するための処理のためにサーバ414などに通信される。

40

【0080】

本実施形態において、図8がソース画像420から決定された顔属性を提示するための顔分析のためのインターフェース（例えばスクリーン800）を示す。本インターフェー

50

スでは、ソース画像が顔属性に関する情報 802 を提示する（例えばオーバーレイする）ように修正された（420A）ものとして提示される。顔属性情報 803 はコンテキスト的に、かつ属性に関連する顔 420 の領域に関連して提示される。本実施形態では、顔属性情報がソース画像 420 について決定された、表 1 からの属性および関連値を識別する。本明細書の画面における、オーバーレイスタイル及び制御位置は例示である。

【0081】

本実施形態では、領域が例えば画像処理によって、例えば領域を特定するために特徴検出を実行することによって、ソース画像 420 上に配置される。ロケーションは例えばデバイス 402 又はサーバ 414 によって決定される。領域は例えば顔全体、眉、目および唇に関連する。これらの領域は、1つ以上の特定の顔属性に関連する。実施形態では、各領域がそれぞれの製品タイプに関連する。

10

【0082】

本実施形態では、個々の情報（例えば 802A）は関連する領域インターフェースに進むための制御（例えばタップ制御 802B）に関連付けられる。各関連領域インターフェースは図示のように、領域の顔属性および関連製品などに関する追加情報を提供するように構成される。図 9A（制御 802B に関連する）および図 9B（制御 802C に関連する）を参照されたい。一実施形態では、各領域がそれぞれの領域界面を有する。

【0083】

それぞれの（タップ）コントロール 804 はまた、顔属性および関連製品についての追加情報を提供する関連領域インターフェースに進むために提供される。この実施形態では、制御 804 がスワイプジェスチャ（例えば左/右）および/または前進制御（例えば 804A）のタッチなどによって前進される（呼び出される）。

20

【0084】

本実施形態では、スクロールがさらなる情報（例えば「私の結果概要」部分）をもたらし、インターフェース 800 の部分 800A を制御する。この実施形態では保存制御 806（例えば「結果画像の保存」）はアノテーション付き画像 420 及び/又は顔属性分析の結果を装置 402 の記憶装置に保存するために提供される。制御 808（例えば「私のメイクアップルックを続ける」）は推奨インターフェース（例えば図 10A）に進む。コントロールなどの様々なインターフェース部分および要素は、インターフェースナビゲーション及びユーザインターフェース理解を支援するためにテキストラベルを有することができることが明らかである。

30

【0085】

本実施形態において、図 9A が顔全体の顔属性に関連する領域 902 を視覚的に強調し、顔の形状、肌の色合い、アンダートーン（例えば「形、肌の色&アンダートーン情報メイクアップチップ」）などのさらなる情報（例えば部分 904 における）を提供するための領域インターフェース 900 を示す。部分 904 における追加情報は領域 902 の関連製品に関し、グラフィックス及び/又はテキスト、カラスウォッチ画像等を含む。この実施形態では、製品推奨インターフェース（例えば図 10A）に進むために、制御 906 が提供される（例えば「私の推薦を見る」）。制御部 908 は別の領域インターフェース、例えば眉（例えば図 9B）に進むために設けられる。

40

【0086】

本実施形態において、図 9B が領域インターフェース 910 を示し、額の顔属性に関連する領域 912 を視覚的に強調し、額の形状、額の色（例えば「額の形&色情報メイクアップチップ」）などの更なる情報を部分 914 に提供する。部分 914 における追加情報は領域 912 の関連製品に関し、グラフィックス及び/又はテキスト、カラスウォッチ画像等を含む。製品推奨インターフェース（例えば図 10A）に進むために、制御 906 が提供される（例えば「私の推薦を見る」）。コントロール 918 は、目のためなどの別の領域インターフェース（目のための領域インターフェースは図示せず）に進むために提供される。コントロール 920 は特定の領域インターフェースまたは「私の結果」インターフェース（例えば図 8）に進むために提供される。特定の領域インターフェースに進む

50

ことは、例えばコントロール 9 2 0 のうちの 1 つをスワイプするか、又はテーピングすることによってナビゲートされ得る。

【 0 0 8 7 】

領域界面（例えば 9 0 0 及び 9 1 0）について、関心領域（9 0 2, 9 1 2）はソース画像 4 2 0 上に（例えばオーバーレイを介して）位置する。画像処理は例えば領域を特定するために特徴検出を実行する。本実施形態では、ソース画像 4 2 0 をオーバーレイしてそれぞれの修正画像 4 2 0 B, 4 2 0 C を定義するように、それぞれのマスクが定義される。本実施形態では、マスクが例えばそれぞれの領域に焦点を合わせるために、関心領域（9 0 2、9 1 2）の外側の領域を暗くする。

【 0 0 8 8 】

図示されていないが、本実施形態では唇および目のための領域インターフェースが設けられている。唇領域インターフェースはリップサイズ情報を提示し、唇形状情報を示し、かつ/または記述する。目領域インターフェースは、目の形状、まぶた及び目の色情報を提示する。図示された領域インターフェースの全てと同様に、（例えば領域に関連付けられた適用可能な顔属性に関連する）製品チップも提供される。

【 0 0 8 9 】

本実施形態では、図 1 0 A ~ 1 0 D が推奨インターフェースの初期画面 1 0 0 0、1 0 1 0、1 0 2 0 及び 1 0 3 0 を示す。画面 1 0 0 0、1 0 1 0、1 0 2 0 および 1 0 3 0 は様々な製品情報を提示し、前進制御（例えば「次へ」1 0 0 2）を含む。スクリーン 1 0 0 0 は領域またはアイテム情報に進むためのコントロール 1 0 0 4 A 及び 1 0 0 4 B を有する「パーソナルルック導入画面」を提供し、全体ルック情報をそれぞれ提供する。図 1 0 B のルック選択インターフェース 1 0 1 0 は「ルック 1」、「ルック 2」、「ルック 3」及び「ルック 4」（例えば全体的な顔の影響を達成するためにそれぞれの顔の影響を有する所定の又は選択されたメイクアップ製品グループ）などの所定の「ルック」に対する選択肢（例えば選択コントロール 1 0 1 2 A、1 0 1 2 B、1 0 1 2 C 及び 1 0 1 2 D に関連する）を提示する。図では一般的にラベル付けされているが、各ルックは描写的又は空想的なラベル（例えば「ドラマチック」、「エレガント」、「ビジネス」、「ランニング・レイト」など）を有することができ、メイクアップタイプは各（顔）領域に関連付けられ、インターフェースを操作し、追加の情報提示を要求するために選択可能（例えば呼び出し可能）である。情報は、例えばデータベース 4 1 6 に記憶される。この実施形態では、外観がそれを適用する製品または技法が 1 又は複数の特定の顔属性に関連付けられる（例えば推奨される）など、顔属性情報にキー付けされる（関連付けられて記憶される）。例えば楕円形の顔、アーモンドの目、厚い眉、ブラウンの眉、赤色のアンダートーン、ブロンドの髪等を有する顔の特定の外観を達成するために、顔、目、眉および唇に適用可能な製品が決定される。

【 0 0 9 0 】

本実施形態では、肌のタイプ選択インターフェース（例えば画面 1 0 2 0）は「タイプ 1」、「タイプ 2」、「タイプ 3」及び「タイプ 4」などのスキントップ選択のための選択肢（例えば選択コントロール 1 0 2 2 A、1 0 2 2 B、1 0 2 2 C 及び 1 0 2 2 D に関連する）を提示する。図では一般的にラベル付けされているが、各タイプは描写的または空想的なラベルを有することができる。本実施形態では、肌のタイプが正常、乾燥、油性、これらの組み合わせ等の属性値に関する。本実施形態では、肌のタイプが画像処理により自動的に決定されないユーザによって提供されるデータである。

【 0 0 9 1 】

本実施形態では、製品推奨が肌のタイプに応答する。次へ 1 0 0 2 を選択することは、アプリケーション 4 0 2 を呼び出し、ルック及び肌のタイプ情報を要求するサーバ 4 1 2 へのメッセージ等により製品推奨を決定するために使用する、ルック及び肌のタイプ情報を提供する。本実施形態では、サービスへのメッセージがソース画像 4 2 0 について決定されたユーザの顔属性情報を提供する。一実施形態では、ソース画像 4 2 0 が効果の適用に使用するため等に提供される。ソース画像 4 2 0 を示す画面 1 0 3 0（図 1 0 D）は製

10

20

30

40

50

品推奨を受信するのを待つ場合等に（例えばラベル「ルックマッチ」を用いて）提示される。画面1030は、図11Aの画面100においてシミュレートされた製品推奨と比較するためのベースライン（例えばメイクアップ基準なし）を提供する。

【0092】

本実施形態において、図11A及び図11Bは現実がシミュレートされる製品推奨インターフェース（例えば1100A及び1100Bとして2つの部分に示される画面1100）を示す。顔属性に一致すると決定された製品推奨を含む選択されたルックは、ソース画像420上でシミュレートされ、画像420Dを提示する。

【0093】

本実施形態では、画面1100が所望の（選択された）外観（例えば1102）を達成するための製品のマッチングに使用される顔属性情報（例えば1102）を提示する。この実施形態では、1002の情報がユーザのそれぞれの属性（例えばユーザ詳細1及びユーザ詳細2）のうち2つにとって、ルックが理想的であることを示す。本実施形態では、特定のルックについての一般的な情報（例えば1104）が提供される。一般的な情報はこの実施形態ではルックの名前および特徴を含み、ルックの製品およびチュートリアル情報を含む。共有/保存コントロール1106は、シミュレートされた画像420Dを保存または共有するために提供される。

10

【0094】

本実施形態では、画面1100が、部分1110、1112、1114及び1116における顔の領域（例えば全体的な顔、目、眉および唇）の各々についての推奨情報を提示するために分割される。目のための部分1112は簡潔にするために、部分的にのみ表示されているが、他の部分1110、1114及び1116と同様に構成されていることに留意されたい。部分1110、1114及び1116は特定の製品推奨（例えば1110A、1114A及び1116Aにおける製品画像、色、名前、価格を含む）、ならびに1110B、1114B及び1116Bにおける特定の製品チュートリアルを示す。本実施形態では、各それぞれの製品が製品を商品としてショッピングカートに追加するために、それぞれの購入コントロール（例えば1110C、1114C及び1116Cで「カートに追加」）に関連付けられる。提示すべき情報の量に応答して、一例として、部分1110、1112、1114及び1116のいずれかは、（例えば左または右へのスワイプジェスチャを介して）スクロール可能であり、例えば領域のための2つ以上の製品、チュートリアル及び購入コントロールを提示する。ステップカウント及び/又はボタン表示（例えば1110D、1114D及び1116D）は、画面1100上に現在提示されているそのような情報内の位置を示すのに有用である。

20

30

【0095】

本実施形態では、コントロール1118の操作により全ての製品がショッピングカートに追加される。例えば製品属性に回答して、個々の製品は選択可能な特徴を有する（例えばサイズ又は任意の他の選択可能な属性を選択するために）。部分1110、1112、1114及び1116のいずれかにおける、それぞれの製品のための関連するコントロール（図示せず）を選択することにより、インターフェースは、製品選択インターフェース1120を提示する。例えば部分1110において、製品推奨1110Aに関連するコントロールを選択することで、製品詳細選択1122が推奨1110Aを置き換え、詳細選択1122を使用して選択を確認する選択コントロール1124がカートに追加コントロール1110Cを置き換える、その製品推奨のための1120に類似する適用可能な製品セレクトインターフェースを呼び出す。

40

【0096】

コントロール1126（例えば「私のルックの選択」）は様々なルック間で、又は様々なルック特徴もしくはコンテキストのためのようなルック内で移動するように動作可能である。本実施形態では、コンテキストが時刻、イベント、場所などに関する。呼び出しコントロール1126は、それぞれ、関連する顔効果の新しい製品推奨および新しい現実シミュレーションをもたらす。検索コントロール（1128）（例えば「メイクアップファ

50

ミリーの検索」)は、追加の製品(インターフェースは図示せず)を示す。

【0097】

本実施形態では、画面1110が、ルック/変更の選択を変更する(例えばルックインターフェースをナビゲートする際に1つの段階に戻る)、最初からやり直す(例えば画面700で新しい画像キャプチャを開始する)、結果に戻る(例えば画面800)、終了する(インターフェースを閉じる、図示せず)などの推奨インターフェースのための様々なナビゲーション制御1130A、1130Bを提示する。

【0098】

一実施形態ではプライバシーに関連する懸念のために、412又は414等のサーバはソース画像またはソース画像を処理した後等の任意の導出された画像を永続的に記憶しない。本実施形態では、アプリケーション404が顔属性処理および製品シミュレーションの要求に(関連付けて)ソース画像を送信する。

10

【0099】

顔属性が決定されると、ユーザエクスペリエンスを支援し、適用可能なユーザインターフェースの動作を定義するために、電子商取引サービスにおけるいくつかの目的に有用である。例えばアプリケーション及びそのインターフェースは、1又は複数を実行するように構成可能である。1.ユーザがファンデーションの色合いを間違っている場合は、肌検出を使用して仮想な試着にヒントを追加する。2.髪の色検出を使用して、達成可能なカラーを推奨する。3.髪の色検出を使用して、トレンドイメージから髪の色を抽出し、関連するカラーを推奨する。4.より多くの目のパラメータを使用して「完全な目の外観」を行う。5.共有特性に基づいてユーザに適合する類似の外観を見つける。6.顔/髪の特徴を使用して、ライフスタイル、個人的ファッション嗜好または他のユーザ特徴を予測する。7.ユーザがルックを決定した後、彼らの顔/髪の特徴に基づいて達成するためのルーチン/アクションを提供する。8.魅力度の評価を計算する。

20

【0100】

図12は、一実施形態による動作1200を示すフローチャートである。動作1200は、例えば本明細書で説明される実施形態に従って、訓練および訓練データセットに関連する方法を提供する。ステップ1202において、動作は推論時間画像(inference time image)から属性を予測する属性分類器を訓練するための属性データセットを記憶し、属性データセットは複数の属性を示す複数の画像を備え、属性の各々は複数のそれぞれの属性値を有する。ステップ1204において、動作は複数の画像の各画像のデータセットにソフトラベルを格納し、ソフトラベルは独立して動作する複数のそれぞれの人間画像アノテータ(human image annotators)によって決定されるように、属性ごとにそれぞれの属性値を含む。ステップ1206において、動作は、属性分類器を訓練するための属性データセットを提供する。記憶動作は一実施形態では限定はされないが、データベース等のデータ記憶装置に記憶する。

30

【0101】

一実施形態では、動作1200が(例えばステップ1208において)属性データセットを使用して属性分類器(attribute classifier)を訓練することをさらに含むことができる。一実施形態では訓練するとき、方法はソフトラベルの中から真を選択するために「真」解決ルールを使用することを含む。

40

【0102】

一実施形態では、属性が属性の固有の関係および異種性に従って複数のそれぞれの属性グループに関連付けられ、属性分類器は属性グループの各々について1つずつ、複数のサブモデルを備える。一実施形態では、それぞれの属性グループのうちの1つは色ベースの属性のための色属性グループである。一実施形態では、それぞれの属性グループのうちの1つは形状ベースの属性のための形状属性グループである。

【0103】

一実施形態では、属性は顔属性である。一実施形態では、顔属性が表1の属性を含む。

【0104】

50

動作 1 2 0 0 の特徴および任意の関連する実施形態はそれに応じて構成されるとき、コンピューティングデバイスの態様およびコンピュータプログラム製品の態様に従って提供されることを理解されたい。

【 0 1 0 5 】

図 1 3 は、一実施形態による動作 1 3 0 0 を示すフローチャートである。動作 1 3 0 0 は、例えば顔属性を決定するための方法を提供する。動作 1 3 0 0 はステップ 1 3 0 2 において、ディープラーニング及び教師付き回帰を実行するネットワークモデルを使用して顔のソース画像を処理し、顔属性の各々について予測を出力する。一実施形態では、ネットワークモデルがディープラーニングを実行して、顔属性を予測するためにそれぞれの分類器による分類のための共有特徴の特徴ベクトルを生成する残差ブロックを備える畳み込みニューラルネットワーク (CNN) モデルを備える。

10

【 0 1 0 6 】

一実施形態では、顔属性が顔属性の固有の関係および異質性に従って複数のそれぞれの属性グループに関連付けられ、ソース画像は属性グループのそれぞれの 1 つのために構成されたそれぞれのネットワークモデルによって処理される。一実施形態ではそれぞれの属性グループのうちの一つは色ベースの顔属性のための色属性グループであり、ソース画像は色ベースのネットワークモデルによって処理される。一実施形態ではそれぞれの属性グループのうちの一つは形状ベースの顔属性のための形状属性グループであり、ソース画像は形状ベースのネットワークモデルによって処理される。

【 0 1 0 7 】

一実施形態では、動作 1 3 0 0 がステップ 1 3 0 4 において、顔属性のそれぞれの 1 つについて、それぞれの分類器によって処理されるべき特徴ベクトルを複製する。

20

【 0 1 0 8 】

一実施形態では、ネットワークモデルが複数のそれぞれの分類器を備え、それぞれの分類器の各々は 1 又は複数の完全に接続された線形レイヤを備え、それぞれの分類器の各々は顔属性のそれぞれ 1 つの予測を出力として提供する。一実施形態では、複数のそれぞれの分類器が顔属性を提供するために並行して実行する。

【 0 1 0 9 】

一実施形態では、モデルが予測レイヤのトリミングされた、適応され事前訓練された ResNet ベースの画像処理ネットワークモデルを含む。

30

【 0 1 1 0 】

一実施形態では、ステップ 1 3 0 6 において、動作は画像検索、顔認識ならびに製品および/またはサービス推奨のうちの一つを実行するための (例えばアプリケーションによる使用のための) 顔属性を提供する。

【 0 1 1 1 】

一実施形態では、ステップ 1 3 0 8 において、動作は画像検索、顔認識ならびに製品および/またはサービス推奨のうちの一つを実行する。

【 0 1 1 2 】

一実施形態では、アプリケーションが顔属性に回答して製品および/またはサービス推奨を実行し、アプリケーションは推奨される製品またはサービスに関連する効果をシミュレートするようにソース画像を修正し、ユーザエクスペリエンスに仮想的な試着を提供する。

40

【 0 1 1 3 】

動作 1 3 0 0 の特徴および任意の関連する実施形態は、それに応じて構成されるとき、コンピューティングデバイスの態様およびコンピュータプログラム製品の態様に従って提供されることを理解されたい。

【 0 1 1 4 】

図 1 4 は、一実施形態による動作 1 4 0 0 を示すフローチャートである。動作 1 4 0 0 は、顔画像の顔属性を決定する方法を提供する。ステップ 1 4 0 2 において、動作は、残差ブロックを有する畳み込みニューラルネットワーク (CNN) ベースのバックボーンネ

50

ットワークモデルを使用して顔のソース画像を処理して、共有特徴の特徴ベクトルを生成する。ステップ1402では、動作が、複数（K個）の分類器を使用して特徴ベクトルを処理し、K個のそれぞれの顔属性が決定され、それぞれの分類器は、K個のそれぞれの顔属性のうち1つの予測を出力するために1又は複数の完全に接続された線形レイヤと予測ブロックとを備える。

【0115】

一実施形態によれば、処理ステップ（1402，1404）は、顔のN個のソース画像に対して実行され、予測解決ルールを使用して、各顔属性のN個の予測からK個のそれぞれの顔属性の最終予測を決定する。

【0116】

一実施形態によれば、CNNベースのバックボーンネットワークモデルは、第1のネットワークモデルと第2のネットワークモデルとを含む。一実施形態によれば、第1のネットワークモデルは、色ベースの顔属性についてソース画像を処理して、色ベースの顔属性を予測するように構成された複数の分類器のそれぞれによる処理のための色ベースの特徴ベクトルを生成する。

【0117】

一実施形態によれば、第2のネットワークモデルは、形状ベースの顔属性についてソース画像を処理して、形状ベースの顔属性を予測するように構成された複数の分類器のそれぞれによる処理のための形状ベースの特徴ベクトルを生成する。

【0118】

一実施形態によれば、特徴ベクトルは、K個の分類器による並列処理のために複製される。

【0119】

一実施形態では、ステップ1406において、動作は画像検索、顔認識、ならびに製品および/またはサービス推奨のうちいずれか1つを実行するための（例えばアプリケーションによる使用のための）顔属性を提供する。

【0120】

一実施形態では、ステップ1408において、動作は画像検索、顔認識ならびに製品および/またはサービス推奨のうちいずれか1つを実行する。

【0121】

一実施形態では、アプリケーションが顔属性に回答して製品および/またはサービス推奨を実行し、アプリケーションは推奨される製品またはサービスに関連する効果をシミュレートするようにソース画像を修正し、ユーザエクスペリエンスに仮想的な試着を提供する。

【0122】

動作1400の特徴および任意の関連する実施形態はそれに応じて構成されるとき、コンピューティングデバイスの態様およびコンピュータプログラム製品の態様に従って提供されることを理解されたい。

【0123】

図15は、一実施形態による動作1500を示すフローチャートである。動作1500は、現実をシミュレートするための方法を提供する。1502において、動作は、顔のソース画像から複数の顔属性を決定し、顔属性分類ネットワークモデルを使用してソース画像を処理する。1504において、動作は、顔属性のうち少なくとも1つに回答して、顔に適用するための少なくとも1つの顔効果を決定する。動作（1506）は、ソース画像を処理して、少なくとも1つの顔効果を表示のために顔に適用する。

【0124】

一実施形態では、動作1500の顔属性が、そのようなそれぞれの動作1200又は1300の任意の関連する実施形態を含む、それぞれの動作1200又は1300を使用して決定される。動作1500は、従って、任意の関連する実施形態を含むそのような動作1200又は1300を実行することを含み得る。

10

20

30

40

50

【0125】

一実施形態では、少なくとも1つの顔効果が顔面に適用される少なくとも1つのメイクアップ製品および/または技法のシミュレーションである。

【0126】

一実施形態では、少なくとも1つの顔効果を適用するためにソース画像を処理することは、顔効果を適用するように構成されたディープラーニングネットワークを使用することを含む。

【0127】

一実施形態では、動作1500のステップ1508が、顔属性に関連するメイクアップ製品および/または技法を記憶するデータストアから、1又は複数のメイクアップ製品および/または技法を選択するために、顔属性のうちの少なくとも1つを使用する。

10

【0128】

動作1500の特徴および任意の関連する実施形態は、それに応じて構成されるとき、コンピューティングデバイスの態様およびコンピュータプログラム製品の態様に従って提供されることを理解されたい。

【0129】

図16は、一実施形態による動作1600を示すフローチャートである。動作1600は、電子商取引サービスのために製品を推奨する方法を提供する。ステップ1602において動作は、顔のソース画像から決定された複数の顔属性を受信し、ソース画像は、複数の顔属性を生成するために顔属性分類ネットワークモデルを使用して処理される。1604において動作は、製品に適した顔属性に関連した製品を記憶するデータストアから、少なくとも1つの製品を選択するために、少なくともいくつかの顔属性を使用する。1606において動作は、少なくとも1つの製品を、製品を購入するための電子商取引インターフェースにおける提示のための推奨として提供する。

20

【0130】

一実施形態では、動作1600の顔属性がそのようなそれぞれの動作1200又は1300の任意の関連する実施形態を含むそれぞれの動作1200又は1300を使用して決定される。動作1600は、従って、任意の関連する実施形態を含むそのような動作1200又は1300を実行することを含み得る。

【0131】

一実施形態では、製品がメイクアップ製品を含む。

30

【0132】

一実施形態では、メイクアップ製品が所定のメイクアップルックを定義するためにデータストア内で様々に関連付けられ、動作(図示せず)は所定のメイクアップルックのうちの1つの識別を受信し、顔属性のうちの少なくともいくつかを使用するステップは、少なくとも1つの製品を選択するとき、所定のメイクアップルックのうちの1つに回答する。一実施形態では、メイクアップ製品の各々が複数のメイクアップタイプのうちの1つに関連付けられ、方法は推奨を定義するために、メイクアップタイプの各々について、顔属性に回答して少なくとも1つの製品を選択することを含む。一実施形態では、メイクアップタイプが、顔製品タイプ、目製品タイプ、眉製品タイプ及び唇製品タイプを含む。

40

【0133】

一実施形態では、動作(図示せず)が推奨製品を使用する技術を更に推奨する。

【0134】

一実施形態では例えば電子商取引インターフェースを含むことができるインターフェースを介して、動作(図示せず)は、ソース画像を処理することによって顔に適用される少なくとも1つの製品のシミュレーションを提供する。

【0135】

一実施形態では、動作(図示せず)は、推奨製品の少なくともいくつかを購入するための電子商取引ショッピングサービスを提供する。

【0136】

50

動作 1 6 0 0 の特徴および任意の関連する実施形態はそれに応じて構成されるとき、コンピューティングデバイスの態様およびコンピュータプログラム製品の態様に従って提供されることを理解されたい。

【 0 1 3 7 】

本明細書に示され、説明されるコンピューティングデバイス及びインターフェースは、複数の異なる態様を提供する。例えば一実施形態ではコンピューティングデバイスがシステムの構成要素など、構成可能であり、システムはそれぞれの機能を実行するように構成された処理回路を含む 1 又は複数の特定の機能ユニットを備える。

【 0 1 3 8 】

1 つのそのような態様では、図 1 7 に示すように、顔のソース画像から複数の顔属性を抽出し、抽出した複数の顔属性に基づいて 1 又は複数の顔効果を生成するように構成された処理回路を含む顔属性ユニット 1 7 0 2 と、少なくとも 1 つの顔効果をソース画像に適用し、効果が適用されたソース画像の 1 又は複数の仮想インスタンスを電子商取引インターフェース上で生成するように構成された処理回路を含む顔効果ユニット 1 7 0 4 とを含むシステム 1 7 0 0 が提供される。

10

【 0 1 3 9 】

一実施形態では、システムが複数の顔属性のうち 1 又は複数に関連して製品を記憶するデータストアから少なくとも 1 つの製品を決定し、製品を購入するために電子商取引インターフェース上で製品推奨の 1 又は複数の仮想インスタンスを生成するように構成された処理回路を含むユーザエクスペリエンスユニット 1 7 0 6 をさらに備える。

20

【 0 1 4 0 】

一実施形態では、顔属性ユニットがディープラーニング及び教師付き回帰を実行して、顔のソース画像から複数の顔属性の各々についての予測を出力するように構成されたネットワークモデルを実行するように構成された処理回路を含む。

【 0 1 4 1 】

一実施形態では、顔効果ユニットは、効果が適用されたソース画像の少なくとも一部と、メイクアップ製品またはメイクアップ適用技術のうち少なくとも 1 つとを含む仮想表現 (virtual representation) を生成するように構成された処理回路を含む。

【 0 1 4 2 】

一実施形態では、顔効果が顔属性のうち 1 つに応答するアノテーションを含む。一実施形態ではアノテーションが顔属性の顔属性値を含み、及び / 又はアノテーションが顔属性が決定されるソース画像の領域に位置する (例えばソース画像上に (例えばオーバーレイとして) 位置する) 。

30

【 0 1 4 3 】

一実施形態では、顔属性ユニット 1 7 0 2 によって決定される顔属性が、動作 1 2 0 0 又は 1 3 0 0 それぞれの任意の関連する実施形態を含む、動作 1 2 0 0 又は 1 3 0 0 それぞれを使用して決定される。従って、顔属性ユニットは、任意の関連する実施形態を含む動作 1 2 0 0 又は 1 3 0 0 を実行するように構成され得る。

【 0 1 4 4 】

他の実施形態では、他の実用的な用途が本明細書に記載の顔属性分類器を使用する。他の実施形態は、画像検索、顔認識などを含む。画像検索の主なタスクは、非常に大きなデータベースから画像を検索することである (この大きなサイズでは、順次検索は通常機能しない) 。一実施形態では、顔属性が画像データベースに記憶された画像について計算され、それに応じてキー入力される。検索のための候補画像もまた、計算された属性を有し、これらは検索基準を狭くするため、又は、信頼度一致データ (confidence match data) を定義するため、又は、検索結果を順序付けるために使用される (例えば候補画像は 1 つの特定の記憶された画像の 7 つの顔属性のうち 5 つと、第 2 の特定の記憶された画像の 7 つのうち 6 つと一致し、その結果は検索結果を順序付けるため、又は (部分的に) 一致する記憶された画像のうち 1 つを選択するために使用される) 。

40

【 0 1 4 5 】

50

例えば ResNet のような CNN モデルは画像を「特徴」として表すことができ、これは、元の画像の圧縮された表現であり、従って、通常、はるかに小さく、検索がより容易である。

【0146】

顔認識では、そのようなタスクが主に、同じアイデンティティ（人物）の異なる画像を見つけること、分類すること、又は、グループ化することに関する。同様の技法は、顔属性を使用して実行可能である。

【0147】

本開示は、以前の著作物またはデータセットにおいて以前にカバーされなかった（顔の）属性の特別なセットを予測する方法を可能にする、包括的なデータセットのコレクションを説明する。以前の研究（例えば「CelebFaces Attributes dataset」（CelebA）データセット（URL: mmlab.ie.cuhk.edu.hk/projects/CelebA.html）で入手可能な香港の中国大学のマルチメディア研究所からの顔属性データセット）は畳み込みニューラルネットワーク（CNN）を以前に探索したが、そのようなものは各属性に対して異なるブランチを設定しなかった。即ち異なる属性ごとに別個の分類器ヘッドは使用されなかった。色および形状などの固有の関係を有する異なる属性に対して、異なるサブモデルを使用しなかった。

本明細書に示され、説明されるようなデータセットの収集の実施形態では、訓練ワークフローがモデル精度を改善するためにソフトラベルを使用した。訓練から得られる属性分類子は、人間のアノテーションと同等である。

【0148】

実用的な実装は、本明細書に記載される特徴のいずれかまたは全てを含むことができる。これらおよび他の態様、特徴および様々な組合せは、機能を実行するための方法、機器、系、手段および本明細書で説明する特徴を組み合わせる他の方法として表され得る。いくつかの実施形態について説明した。それにもかかわらず、本明細書に記載されるプロセスおよび技法の趣旨および範囲から逸脱することなく、様々な修正がなされ得ることが理解されよう。加えて、他のステップを提供することができ、又はステップに記載されたプロセスから排除でき、他の構成要素に記載されたシステムに追加するか、又は、そこから除去できる。従って、他の態様は特許請求の範囲の範囲内にある。

【0149】

本明細書の説明および特許請求の範囲を通して、単語「含む（comprise）」及び「含む（contain）」及びそれらの変形は「含むが、限定されない（including but not limited to）」を意味し、他の構成要素、整数またはステップを排除することを意図しない（。本明細書全体を通して、単数形は文脈が他のことを必要としない限り、複数形を包含する。特に、不定冠詞が使用される場合は本明細書がその状況が他のことを要求していない限り、単数だけでなく複数も意図していると理解されたい。

【0150】

本発明の特定の態様、実施形態または実施例に関連して説明される特徴、整数、特性または群はそれらと互換性がない場合を除き、任意の他の態様、実施形態または実施例に適用可能であると理解されるべきである。本明細書に開示される特徴の全て（任意の添付の特許請求の範囲、要約および図面を含む）および/またはそのように開示される任意の方法またはプロセスのステップの全ては、そのような特徴および/またはステップの少なくともいくつかは相互に排他的である組合せを除いて、任意の組合せで組み合わせることができる。本発明は、前述の任意の例または実施形態の詳細に限定されない。本発明は、本明細書（添付の特許請求の範囲、要約および図面を含む）に開示される特徴の任意の新規な1つまたは任意の新規な組み合わせ又は開示される任意の方法またはプロセスのステップの任意の新規な1つまたは任意の新規な組み合わせに及ぶ。

参考文献

1. "Deep Residual Learning for Image Recognition", He, Kaiming et. al, 2015-12-10, available at URL arxiv.org/abs/1512.03385.

10

20

30

40

50

2. A Survey of Deep Facial Attribute Analysis, Xin Zheng, Xin et. al, Submitted on 26 Dec 2018 (v1), last revised 27 Oct 2019 (this version, v3) URL arxiv.org/abs/1812.10265.

3. Deep Learning Face Attributes in the Wild, Liu, Ziwei et. al, Submitted on 28 Nov 2014 (v1), last revised 24 Sep 2015 (this version, v3), URL: arxiv.org/abs/1411.7766.

<その他>

<手段>

技術的思想 1 の方法は、顔の属性を決定する方法であって、ディープラーニング及び教師付き回帰を実行するネットワークモデルを使用して顔のソース画像を処理し、顔属性の各々について予測を出力することと、前記ネットワークモデルは、ディープラーニングを実行する残差ブロックを含む畳み込みニューラルネットワーク（CNN）モデルを含み、前記顔属性を予測するために、それぞれの分類器による分類のための共有特徴の特徴ベクトルを生成することを含む。

10

技術的思想 2 の方法は、技術的思想 1 記載の方法において、前記顔属性は、前記顔属性の固有の関係および異質性に従って複数のそれぞれの属性グループに関連付けられ、前記ソース画像は、前記属性グループのうちのそれぞれの 1 つのために構成されたそれぞれのネットワークモデルによって処理される。

技術的思想 3 の方法は、技術的思想 2 記載の方法において、前記それぞれの属性グループのうちの 1 つは、色ベースの顔属性のための色属性グループであり、前記ソース画像は、色ベースのネットワークモデルによって処理される。

20

技術的思想 4 の方法は、技術的思想 2 又は 3 に記載の方法において、前記それぞれの属性グループのうちの 1 つは、形状ベースの顔属性のための形状属性グループであり、前記ソース画像は、形状ベースのネットワークモデルによって処理される。

技術的思想 5 の方法は、技術的思想 1 から 4 のいずれかに記載の方法において、前記顔属性のそれぞれについて、それぞれの分類器によって処理されるべき前記特徴ベクトルを複製することを含む。

技術的思想 6 の方法は、技術的思想 1 から 5 のいずれかに記載の方法において、前記ネットワークモデルが複数のそれぞれの分類器を備え、前記それぞれの分類器の各々が、1 又は複数の完全に接続された線形レイヤを備え、前記それぞれの分類器の各々が前記顔属性のそれぞれの 1 つの予測を出力として提供する。

30

技術的思想 7 の方法は、技術的思想 6 記載の方法において、前記複数のそれぞれの分類器は、前記顔属性を提供するために並行して実行する。

技術的思想 8 の方法は、技術的思想 1 から 7 のいずれかに記載の方法において、前記モデルは、予測レイヤのトリミングされた、適応され事前訓練された ResNet ベースの画像処理ネットワークモデルを含む。

技術的思想 9 の方法は、技術的思想 1 から 8 のいずれかに記載の方法において、画像検索、顔認識ならびに製品および/またはサービス推薦のうちのいずれか 1 つを実行するための前記顔属性を提供することを含む。

技術的思想 10 の方法は、技術的思想 9 記載の方法において、アプリケーションが前記顔属性に回答して前記製品および/またはサービス推奨を実行し、前記アプリケーションが、推奨される製品またはサービスに関連する効果をシミュレートするように前記ソース画像を修正して、ユーザエクスペリエンスに仮想的な試着を提供する。

40

技術的思想 11 の方法は、顔画像の顔の属性を決定する方法であって、残差ブロックを有する畳み込みニューラルネットワーク（CNN）ベースのバックボーンネットワークモデルを使用して顔のソース画像を処理して、共有特徴の特徴ベクトルを生成することと、複数（K 個）の分類器を使用して特徴ベクトルを処理し、K 個のそれぞれの顔属性が決定され、それぞれの分類器は、K 個のそれぞれの顔属性のうちの 1 つの予測を出力するために 1 又は複数の完全に接続された線形レイヤと予測ブロックとを備えることとを含む。

技術的思想 12 の方法は、技術的思想 11 記載の方法において、前記顔の N 個のソース画

50

像に対して前記処理ステップを実行するステップと、各顔属性の前記N個の予測から前記K個のそれぞれの顔属性の最終予測を決定するために予測解決ルールを使用することを含む。

技術的思想13の方法は、技術的思想11又は12に記載の方法において、第1のネットワークモデルが、色ベースの顔属性について前記ソース画像を処理して、色ベースの顔属性を予測するように構成された前記複数の分類器のそれぞれによって処理するための色ベースの特徴ベクトルを生成する。

技術的思想14の方法は、技術的思想11から13のいずれかに記載の方法において、第2のネットワークモデルが、形状ベースの顔属性について前記ソース画像を処理して、形状ベースの顔属性を予測するように構成された前記複数の分類器のそれぞれによって処理するための形状ベースの特徴ベクトルを生成する。

技術的思想15の方法は、技術的思想11から14のいずれかに記載の方法において、前記K個の分類器による並列処理のために前記特徴ベクトルを複製する。

技術的思想16の方法は、現実をシミュレートする方法であって、顔のソース画像から複数の顔属性を決定し、顔属性分類ネットワークモデルを使用してソース画像を処理することと、前記顔属性のうち少なくとも1つにตอบสนองして前記顔に適用するための少なくとも1つの顔効果を決定することと、表示のために、少なくとも1つの顔効果を顔に適用するために、ソース画像を処理することを含む。

技術的思想17の方法は、技術的思想16に記載の方法において、前記顔属性が、請求項1から15のいずれかに記載の方法を使用して決定される。

技術的思想18の方法は、技術的思想16又は17に記載の方法において、前記少なくとも1つの顔効果は、前記顔に適用される少なくとも1つのメイクアップ製品および/または技法のシミュレーションである。

技術的思想19の方法は、技術的思想16から18のいずれかに記載の方法において、前記少なくとも1つの顔効果を適用するために前記ソース画像を処理することは前記顔効果を適用するように構成されたディープラーニングネットワークを使用することである。

技術的思想20の方法は、技術的思想18又は19に記載の方法において、前記顔属性のうち少なくとも1つを使用して、前記顔属性に関連してメイクアップ製品および/または技法を記憶するデータストアから1又は複数のメイクアップ製品および/または技法を選択することを含む。

技術的思想21の方法は、電子商取引サービスのための製品を推奨する方法であって、顔のソース画像から決定された複数の顔属性を受信し、前記ソース画像は、前記複数の顔属性を生成するために、顔属性分類ネットワークモデルを使用して処理されることと、製品に適した顔属性と関連して製品を記憶するデータストアから少なくとも1つの製品を選択するために、少なくともいくつかの前記顔属性を使用することと、少なくとも1つの製品を、製品を購入するための電子商取引インターフェースにおける提示のための推奨として提供することを含む。

技術的思想22の方法は、技術的思想21に記載の方法において、前記顔属性が、技術的思想1から15のいずれかに記載の方法を使用して決定される。

技術的思想23の方法は、技術的思想21又は22に記載の方法において、前記製品が、メイクアップ製品を含む。

技術的思想24の方法は、技術的思想23に記載の方法において、前記メイクアップ製品は、所定のメイクアップルックを定義するために、データストア内で様々に関連付けられ、当該方法は、所定のメイクアップルックのうち1つの識別を受信することを含み、前記顔属性のうち少なくともいくつかを使用するステップは、少なくとも1つの製品を選択するときに、前記所定のメイクアップルックのうち1つにตอบสนองするものである。

技術的思想25の方法は、技術的思想24に記載の方法において、前記メイクアップ製品の各々が、複数のメイクアップタイプのうち1つに関連付けられ、当該方法が前記推奨を定義するために、前記メイクアップタイプの各々について、前記顔属性にตอบสนองして、少なくとも1つの製品を選択することを含む。

10

20

30

40

50

技術的思想 26 の方法は、技術的思想 24 記載の方法において、前記メイクアップタイプが、顔製品タイプ、目製品タイプ、眉製品タイプ及び唇製品タイプを含む。

技術的思想 27 の方法は、技術的思想 21 から 26 のいずれかに記載の方法において、推奨製品を使用するための技術を更に推奨する。

技術的思想 28 の方法は、技術的思想 21 から 27 のいずれかに記載の方法において、前記電子商取引インターフェースが、前記ソース画像を処理することによって前記顔に適用される前記少なくとも 1 つの製品のシミュレーションを提供する。

技術的思想 29 の方法は、技術的思想 21 から 28 のいずれかに記載の方法において、前記推奨製品のうちの少なくともいくつかを購入するための電子商取引ショッピングサービスを提供する。

10

技術的思想 30 のシステムは、技術的思想 1 から 28 記載のいずれかの方法を実行するように構成されたコンピューティングデバイスを備える。

技術的思想 31 のシステムは、顔のソース画像から複数の顔属性を抽出し、抽出された複数の顔属性に基づいて 1 又は複数の顔効果を生成するように構成された処理回路を含む顔属性ユニットと、少なくとも 1 つの顔効果を前記ソース画像に適用し、効果が適用されたソース画像の 1 又は複数の仮想インスタンスを電子商取引インターフェース上で生成するように構成された処理回路を含む顔効果ユニットとを含む。

技術的思想 32 のシステムは、技術的思想 31 記載のシステムにおいて、複数の顔属性のうちの 1 又は複数に関連して製品を記憶するデータストアから少なくとも 1 つの製品を決定し、製品を購入するために電子商取引インターフェース上で製品推奨の 1 又は複数の仮想インスタンスを生成するように構成された処理回路を含むユーザエクスペリエンスユニットを更に含む。

20

技術的思想 33 のシステムは、技術的思想 31 又は 32 に記載のシステムにおいて、前記顔属性ユニットは、ディープラーニング及び教師付き回帰を実行するネットワークモデルを実行して、顔の前記ソース画像から前記複数の顔属性の各々についての予測を出力するように構成された処理回路を含む。

技術的思想 34 のシステムは、技術的思想 31 から 33 のいずれかに記載のシステムにおいて、前記顔効果ユニットは、前記効果が適用されたソース画像の少なくとも一部と、メイクアップ製品またはメイクアップ適用技術のうちの少なくとも 1 つとを含む仮想表現を生成するように構成された処理回路を含む。

30

技術的思想 35 のシステムは、技術的思想 31 から 33 のいずれかに記載のシステムにおいて、前記顔効果は、前記顔属性のうちの 1 つに応答するアノテーションを含む。

技術的思想 36 のシステムは、技術的思想 35 記載のシステムにおいて、前記アノテーションは、前記顔属性の顔属性値を含み、前記アノテーションは、前記顔属性が決定される前記ソース画像の領域に位置する。

技術的思想 37 の方法は、推論時間画像から属性を予測する属性分類器を訓練するための属性データセットを記憶すること、前記属性データセットは複数の属性を示す複数の画像を含み、前記属性の各々は複数のそれぞれの属性値を有するものであり、前記複数の画像の各々の画像に対する前記データセットにソフトラベルを記憶すること、前記ソフトラベルは独立して動作する複数のそれぞれの人間画像アノテータによって決定される、属性ごとのそれぞれの属性値を含むものであり、前記属性分類器を訓練するための前記属性データセットを提供することを含む。

40

技術的思想 38 の方法は、技術的思想 37 記載の方法において、前記属性データセットを使用して前記属性分類器を訓練することを更に含む。

技術的思想 39 の方法は、技術的思想 38 記載の方法において、訓練時に、前記ソフトラベルの中から真を選択するために「真」解決ルールを使用することを含む。

技術的思想 40 の方法は、技術的思想 37 から 39 のいずれかに記載の方法において、前記属性は前記属性の固有の関係および異質性に従って複数のそれぞれの属性グループに関連付けられ、前記属性分類器は前記属性グループの各々について 1 つずつ、サブモデルを備える。

50

技術的思想 4 1 の方法は、技術的思想 4 0 記載の方法において、前記それぞれの属性グループのうちの一つは、色ベースの属性のための色属性グループである。

技術的思想 4 2 の方法は、技術的思想 4 0 又は 4 1 に記載の方法において、前記それぞれの属性グループのうちの一つは、形状ベースの属性のための形状属性グループである。

技術的思想 4 3 の方法は、技術的思想 3 7 から 4 2 のいずれかに記載の方法において、前記属性が顔属性である。

技術的思想 4 4 の方法は、技術的思想 4 3 記載の方法において、前記顔属性は、表 1 の属性を含む。

技術的思想 4 5 のシステムは、技術的思想 3 7 から 4 3 のいずれかに記載の方法を実行するように構成されたコンピューティングデバイスを備える。

技術的思想 4 6 のシステムは、技術的思想 3 7 から 4 3 のいずれかに記載の方法に従って定義された属性分類器を用いて構成されたコンピューティングデバイスを備える。

10

20

30

40

50