

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7410735号  
(P7410735)

(45)発行日 令和6年1月10日(2024.1.10)

(24)登録日 令和5年12月26日(2023.12.26)

(51)国際特許分類

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 6 6 0 A

G 0 6 F 3/01 (2006.01)

G 0 6 T 7/00 3 5 0 B

G 0 6 F 3/01 5 1 0

請求項の数 13 (全21頁)

(21)出願番号	特願2020-18213(P2020-18213)	(73)特許権者	000001007
(22)出願日	令和2年2月5日(2020.2.5)		キヤノン株式会社
(65)公開番号	特開2021-124977(P2021-124977 A)	(74)代理人	東京都大田区下丸子3丁目30番2号 110003281
(43)公開日	令和3年8月30日(2021.8.30)		弁理士法人大塚国際特許事務所
審査請求日	令和5年2月2日(2023.2.2)	(72)発明者	三河 睦 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
		審査官	小池 正彦

最終頁に続く

(54)【発明の名称】 端末装置、機能提案方法、プログラム及びサーバ装置

(57)【特許請求の範囲】

【請求項1】

サーバ装置と通信する通信手段と、制御手段と、を備える端末装置であって、  
前記制御手段は、  
前記端末装置において起動すべき機能の提案を求める提案要求を、前記通信手段を介して前記サーバ装置へ送信し、  
前記提案要求に応じて前記サーバ装置により決定される提案機能を特定する提案応答を、前記通信手段を介して前記サーバ装置から受信し、  
前記提案応答により特定される前記提案機能を起動し、  
前記提案要求は、前記端末装置のユーザの顔画像データを含み、  
前記提案機能は、前記サーバ装置において前記顔画像データを学習済みモデルに適用することにより決定され、  
前記制御手段は、  
前記提案機能の前記起動の前後の前記ユーザの表情の変化度を、前記ユーザの顔画像に基づいて評価し、  
評価した前記表情の前記変化度が所定の条件を満たす場合に、前記学習済みモデルの再学習を求める学習要求を、前記通信手段を介して前記サーバ装置へ送信する、  
端末装置。

【請求項2】

請求項1に記載の端末装置であって、

前記学習要求は、前記ユーザを識別するユーザ識別情報を含み、  
前記学習済みモデルの前記再学習は、ユーザ固有のモデルを獲得するためにユーザごとに実行される、  
端末装置。

【請求項 3】

請求項 1 又は 2 に記載の端末装置であって、  
前記制御手段は、  
前記ユーザの表情の変化の評価値を前記ユーザの顔画像に基づいて算出するための評価モデルに、前記提案機能の前記起動の前の前記ユーザの第 1 の顔画像を適用することにより、第 1 の評価値を取得し、

10

前記評価モデルに、前記提案機能の前記起動の後の前記ユーザの第 2 の顔画像を適用することにより、第 2 の評価値を取得し、

前記表情の前記変化度は、前記第 1 の評価値から前記第 2 の評価値への変化に相当する、  
端末装置。

【請求項 4】

請求項 3 に記載の端末装置であって、前記所定の条件は、前記第 1 の評価値と前記第 2 の評価値との間の差が第 1 の閾値を上回ることを含む、端末装置。

【請求項 5】

請求項 3 又は 4 に記載の端末装置であって、  
前記評価モデルは、複数の人間の顔画像であって満足度を教師データとしてラベル付けされた当該顔画像を学習用画像として用いた機械学習を通じて予め獲得され、

20

前記所定の条件は、前記第 1 の評価値から前記第 2 の評価値への前記変化が前記満足度が改善したことを示すことを含む、

端末装置。

【請求項 6】

請求項 4 又は 5 に記載の端末装置であって、前記学習済みモデルの前記再学習は、前記提案機能を提案すべき機能の正解とし、前記第 1 の顔画像を学習用画像として用いて実行される、端末装置。

【請求項 7】

請求項 3 乃至 6 のいずれか 1 項に記載の端末装置であって、前記制御手段は、前記評価モデルに前記第 1 の顔画像を適用することにより取得される前記第 1 の評価値が第 2 の閾値を下回る場合に、前記提案要求を前記サーバ装置へ送信する、端末装置。

30

【請求項 8】

請求項 1 乃至 7 のいずれか 1 項に記載の端末装置であって、  
前記提案要求は、前記端末装置において稼働しているアプリケーションを識別するアプリケーション識別情報を含み、  
前記サーバ装置により決定される前記提案機能は、前記端末装置において稼働している前記アプリケーションに依存して異なる、  
端末装置。

【請求項 9】

請求項 1 乃至 8 のいずれか 1 項に記載の端末装置であって、前記ユーザの顔画像を撮像する撮像手段、をさらに備える、端末装置。

40

【請求項 10】

端末装置により行われる機能提案方法であって、  
前記端末装置において起動すべき機能の提案を求める提案要求を、前記端末装置からサーバ装置へ送信することと、  
前記提案要求に応じて前記サーバ装置により決定される提案機能を特定する提案応答を、前記端末装置により前記サーバ装置から受信することと、  
前記提案応答により特定される前記提案機能を前記端末装置により起動することと、  
を含み、

50

前記提案要求は、前記端末装置のユーザの顔画像データを含み、  
前記提案機能は、前記サーバ装置において前記顔画像データを学習済みモデルに適用することにより決定され、  
前記機能提案方法は、  
前記提案機能の前記起動の前後の前記ユーザの表情の変化度を、前記ユーザの顔画像に基づいて評価することと、  
評価した前記表情の前記変化度が所定の条件を満たす場合に、前記顔画像データを用いた前記学習済みモデルの再学習を求める学習要求を、前記端末装置から前記サーバ装置へ送信することと、  
をさらに含む、機能提案方法。

10

【請求項 1 1】

端末装置のプロセッサに、請求項 1 0 に記載の機能提案方法を行わせるためのコンピュータプログラム。

【請求項 1 2】

端末装置と通信する通信手段と、制御手段と、を備えるサーバ装置であって、  
前記制御手段は、  
前記端末装置において起動すべき機能の提案を求める提案要求であって、前記端末装置のユーザの顔画像データを含む前記提案要求を、前記通信手段を介して前記端末装置から受信し、  
前記顔画像データを学習済みモデルに適用することにより、前記端末装置において起動すべき前記機能を決定し、  
決定した前記機能を特定する提案応答を前記通信手段を介して前記端末装置へ送信することにより、前記端末装置に前記機能を起動させ、  
前記顔画像データを用いた前記学習済みモデルの再学習を求める学習要求が前記通信手段を介して前記端末装置から受信された場合に、決定した前記機能を正解とし、前記顔画像データにより示される顔画像を学習用画像として用いて、前記学習済みモデルの前記再学習を実行する、  
サーバ装置。

20

【請求項 1 3】

端末装置であって、  
ユーザを撮像することにより前記ユーザの顔画像を取得する画像取得部と、  
前記端末装置の機能のうち前記顔画像を学習済みモデルに適用することにより決定される機能を起動する制御部と、  
前記機能の前記起動の前後の前記ユーザの表情の変化度が所定の条件を満たす場合に、前記学習済みモデルの再学習を実行する学習部と、  
を備える端末装置。

30

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本開示は、端末装置、機能提案方法、プログラム及びサーバ装置に関する。

40

【背景技術】

【0 0 0 2】

スマートフォンに代表されるユーザ端末は、近年ますます多機能化及び高機能化している。しかし、ユーザ端末の機能を自在に使いこなすことのできるユーザは少ない。特許文献 1 は、特に文字の視認性を高める機能をユーザが知らないケースに着目し、撮像画像に映るユーザの状態を解析して文字の視認性を高める機能を自動的に実行する携帯端末を開示している。特許文献 1 により開示された携帯端末は、例えばユーザの表情の変化として眉をひそめる動作を撮像画像から検出し、その表情の変化に応じて文字の視認性を高める機能を実行する。

【0 0 0 3】

50

特許文献 2 は、人間の表情の変化をより正確に定量的に把握できるようにするために、人間の顔画像から抽出される画像特徴量の変化を、表情の典型的な種別に関連付けて学習する手法を開示している。

【先行技術文献】

【特許文献】

【0004】

【文献】特開 2013 - 109687 号公報

【文献】特開 2018 - 36734 号公報

【発明の概要】

【発明が解決しようとする課題】

10

【0005】

しかしながら、ユーザの表情の変化には個人差がある。そのため、単に既知のモデル又はパターンに従って顔画像を分析するだけでは、表情認識の十分な精度を確保できず、ユーザに適した機能を適切に提案することが難しい。

【0006】

そこで、本開示は、ユーザの表情に基づいてより適切な機能をユーザに提案することを可能にする仕組みを実現することを目的の 1 つとする。

【課題を解決するための手段】

【0007】

ある観点によれば、サーバ装置と通信する通信手段と、制御手段と、を備える端末装置であって、前記制御手段は、前記端末装置において起動すべき機能の提案を求める提案要求を、前記通信手段を介して前記サーバ装置へ送信し、前記提案要求に応じて前記サーバ装置により決定される提案機能を特定する提案応答を、前記通信手段を介して前記サーバ装置から受信し、前記提案応答により特定される前記提案機能を起動し、前記提案要求は、前記端末装置のユーザの顔画像データを含み、前記提案機能は、前記サーバ装置において前記顔画像データを学習済みモデルに適用することにより決定され、前記制御手段は、前記提案機能の前記起動の前後の前記ユーザの表情の変化度を、前記ユーザの顔画像に基づいて評価し、評価した前記表情の前記変化度が所定の条件を満たす場合に、前記学習済みモデルの再学習を求める学習要求を、前記通信手段を介して前記サーバ装置へ送信する、端末装置が提供される。対応する機能提案方法、プログラム及びサーバ装置もまた提供される。

20

【発明の効果】

【0008】

本開示によれば、ユーザの表情に基づいてより適切な機能をユーザに提案することが可能となる。

【図面の簡単な説明】

【0009】

【図 1】一実施形態に係る情報処理システムの概略的な構成の一例を示す模式図。

【図 2】一実施形態に係るユーザ端末及び提案サーバのハードウェア構成の一例を示すブロック図。

40

【図 3】一実施形態に係る提案処理の概略的な流れの一例を示すシーケンス図。

【図 4】一実施形態に係るユーザ端末及び提案サーバの機能面の構成の一例を示すブロック図。

【図 5】2 種類の学習済みモデルの入出力データの例を説明するための説明図。

【図 6】提案機能の起動の承認をユーザに求めるためのユーザインタフェースの一例を示す説明図。

【図 7】一実施形態に係るユーザ端末により実行される処理の流れの一例を示すフローチャート。

【図 8】一実施形態に係る提案サーバにより実行される処理の流れの一例を示すフローチャート。

50

## 【発明を実施するための形態】

## 【0010】

以下、添付図面を参照して実施形態を詳しく説明する。なお、以下の実施形態は特許請求の範囲に係る発明を限定するものではない。実施形態には複数の特徴が記載されているが、これらの複数の特徴の全てが発明に必須のものとは限らず、また、複数の特徴は任意に組み合わせられてもよい。さらに、添付図面においては、同一若しくは同様の構成に同一の参照番号を付し、重複した説明は省略する。

## 【0011】

## &lt;&lt;1. 概要&gt;&gt;

## &lt;1-1. システムの基本的な構成&gt;

図1は、一実施形態に係る情報処理システム10の概略的な構成の一例を示す模式図である。図1を参照すると、情報処理システム10は、ネットワーク20を介して互いに接続されたユーザ端末100及び提案サーバ200を含む。ネットワーク20は、例えばLAN(Local Area Network)、プロバイダネットワーク若しくはインターネット、又はそれらの2つ以上の組合せといった、いかなる種類のネットワークであってもよい。ユーザ端末100は、ユーザにより操作され、様々なアプリケーションを実行する端末装置である。ユーザ端末100は、例えばスマートフォン、携帯電話、PC(Personal Computer)、タブレット、ナビゲーション装置、デジタルカメラ、プリンタ、スキャナ又は複合機といった、いかなる種類の端末装置であってもよい。提案サーバ200は、ユーザ端末100へ起動すべき機能を提案する役割を有するサーバ装置である。提案サーバ200は、例えばコンピュータ又はワークステーションといった、いかなる種類の情報処理装置であってもよい。なお、以下に説明する提案サーバ200の機能は、単一のコンピュータで実現されてもよく、又は複数のコンピュータが互いに連携することにより実現されてもよい。複数のコンピュータが連携する場合には、それらは例えばLAN又はBluetooth(登録商標)といった通信プロトコルを用いて相互に通信する。

## 【0012】

## &lt;1-2. 各装置のハードウェア構成&gt;

## (1) ユーザ端末

図2(A)は、本実施形態に係るユーザ端末100のハードウェア構成の一例を示している。ユーザ端末100は、コントローラ101、ROM102、RAM103、記録媒体104、カメラ105、入力デバイス106、表示デバイス107、マイクロフォン108、スピーカ109、通信I/F111及び非接触I/F112を備える。

## 【0013】

コントローラ101は、後述するコンピュータプログラムを実行することにより、ユーザ端末100の各部を制御する制御手段である。図2(A)にはユーザ端末100が1つのコントローラ101を備える例を示しているが、ユーザ端末100は2つ以上のコントローラを備えてもよい。ROM(Read-Only Memory)102は、不揮発性のメモリである。ROM102は、例えば、ユーザ端末100の基本ソフトウェア(オペレーティングシステム(OS)ともいう)及び様々なアプリケーションプログラムを記憶する。ROM102は、例えば、提案サーバ200と通信して提案サーバ200から起動すべき機能の提案を受け付けるためのコンピュータプログラムを記憶してもよい。RAM(Random Access Memory)103は、所謂メインメモリである。RAM103は、コントローラ101に作業用の一時的な記憶領域を提供する。RAM103は、例えば、表示デバイス107により表示されるべき画像データを一時的に記憶してもよい。記録媒体104は、カメラ105により生成される撮像画像データを記録する媒体である。記録媒体104は、ユーザ端末100に内蔵されていてもよく、又はユーザ端末100に着脱可能であってもよい。

## 【0014】

カメラ105は、例えば、光学系、撮像素子及び画像処理ユニットを含み得る撮像手段である。カメラ105は、被写体の光学的な像を撮像素子での光電変換を通じて電気信号

10

20

30

40

50

へ変換することにより、撮像画像を生成する（即ち、被写体を撮像する）。そして、カメラ105は、例えばノイズ低減といった画像処理を行った後、撮像画像データをコントローラ101へ出力する。撮像画像データは、例えば、コントローラ101により記録媒体104に記録される。なお、本実施形態において、カメラ105は、少なくとも、ユーザ端末100の表示デバイス107の画面と同じ側に向けた画角で被写体を撮像可能な内側カメラを含む。入力デバイス106は、ユーザ端末100を使用するユーザからの指示及び情報入力を受け付けるデバイスである。入力デバイス106は、例えば、ボタン、スイッチ、キーパッド、タッチセンサ、キーボード及びポインティングデバイスのうちの1つ以上を含み得る。表示デバイス107は、ユーザへ提示されるべき画像及び情報を表示するデバイスである。表示デバイス107は、タッチスクリーンを含んでもよい。なお、表示デバイス107の代わりに、何らかの接続インタフェースを介してユーザ端末100と接続される外部のディスプレイが、ユーザ端末100のユーザのために画像又は情報を表示してもよい。マイクロフォン108は、ユーザにより発せられる音声を集音する音声取得手段である。マイクロフォン108は、集音した音声の内容を表す入力音声信号をコントローラ101へ出力する。スピーカ109は、コントローラ101からの出力音声信号に基づいて音声を再生する音声再生手段である。本明細書では、ユーザ端末100によるユーザとのインタラクションが表示デバイス107により表示されるGUI（Graphical User Interface）を介して行われる例を主に説明する。しかしながら、説明されるインタラクションは、マイクロフォン108及びスピーカ109を介する音声インタラクションとして実現されてもよい。

10

20

#### 【0015】

通信インタフェース（I/F）111は、ユーザ端末100による他の装置との通信のための通信手段である。通信I/F111は、例えば、アンテナ、RF（Radio Frequency）回路及びベースバンド回路を含む無線通信インタフェースであってもよい。通信I/F111は、例えば無線LAN、LTE、LTE-Advanced又は3GPP NRといった、いかなる種類の無線通信プロトコルをサポートしてもよい。通信I/F111は、画像データの通信のために、例えば無線LANを利用するPTP/IP（Picture Transfer Protocol over Internet Protocol）を用いてもよい。また、通信I/F111は、NIC（Network Interface Card）のような有線通信インタフェースであってもよい。通信I/F111は、例えばLAN、HDMI（登録商標）又はIEEE1394といった、いかなる種類の有線通信プロトコルをサポートしてもよい。本実施形態において、ユーザ端末100は、通信I/F111を介して提案サーバ200と通信する。ユーザ端末100は、提案サーバ200と直接的に接続してもよく、又はアクセスポイントのような中間的なノードを介して間接的に接続してもよい。非接触I/F112は、ユーザ端末100による非接触通信のための近距離通信手段である。非接触I/F112は、例えばアンテナ、変復調回路及び通信コントローラを含み得る。非接触I/F112は、例えばISO/IEC18092規格に従って動作するNFC（Near Field Communication）インタフェースであってもよい。その代わりに、非接触I/F112は、例えばIEEE802.15.1規格に従って動作するBluetooth（登録商標）インタフェースであってもよい。なお、ユーザ端末100は、USB（Universal Serial Bus）又はIrDA（Infrared Data Association）といった他の種類の通信プロトコルをサポートしてもよい。ユーザ端末100と提案サーバ200との間の通信は、通信I/F111の代わりに非接触I/F112を介して行われてもよい。

30

40

#### 【0016】

##### （2）提案サーバ

図2（B）は、本実施形態に係る提案サーバ200のハードウェア構成の一例を示している。提案サーバ200は、CPU201、ROM202、RAM203、HDD204、通信I/F205、入力デバイス206、表示デバイス207及びGPU208を備える。

#### 【0017】

50

C P U (Central Processing Unit) 2 0 1 は、後述するコンピュータプログラムを実行することにより、提案サーバ 2 0 0 の提案機能を実現するプロセッサである。図 2 ( B ) には提案サーバ 2 0 0 が 1 つの C P U 2 0 1 を備える例を示しているが、提案サーバ 2 0 0 は 2 つ以上の C P U を備えてもよい。R O M 2 0 2 は、不揮発性のメモリである。R O M 2 0 2 は、例えば、提案サーバ 2 0 0 のブートプログラム及び基本ソフトウェアを記憶する。R A M 2 0 3 は、所謂メインメモリである。R A M 2 0 3 は、C P U 2 0 1 に作業用の一時的な記憶領域を提供する。H D D (Hard Disk Drive) 2 0 4 は、所謂補助記憶デバイスである。H D D 2 0 4 は、例えば、ユーザ端末 1 0 0 から受信される提案要求に応じて、ユーザ端末 1 0 0 において起動すべき機能をユーザ端末 1 0 0 へ提案するためのコンピュータプログラムを記憶してもよい。なお、ハードディスクの代わりに例えば半導体メモリ又は光ディスクといった他の種類の記憶媒体が、提案サーバ 2 0 0 の補助記憶デバイスにおいて利用されてもよい。

10

#### 【 0 0 1 8 】

通信 I / F 2 0 5 は、提案サーバ 2 0 0 による他の装置との通信のための通信手段である。通信 I / F 2 0 5 は、例えば、アンテナ、R F 回路及びベースバンド回路を含む無線通信インタフェースであってもよく、又は N I C のような有線通信インタフェースであってもよい。通信 I / F 2 0 5 は、ユーザ端末 1 0 0 の通信 I / F 1 1 1 に関連して上で例示したような、いかなる種類の通信プロトコルをサポートしてもよい。本実施形態において、提案サーバ 2 0 0 は、通信 I / F 2 0 5 を介してユーザ端末 1 0 0 と通信する。

#### 【 0 0 1 9 】

20

入力デバイス 2 0 6 は、提案サーバ 2 0 0 への指示及び情報入力を受け付けるデバイスである。入力デバイス 2 0 6 は、例えば、ボタン、スイッチ、キーパッド、タッチセンサ、キーボード及びポインティングデバイスのうちの 1 つ以上を含み得る。表示デバイス 2 0 7 は、画像及び情報を表示するデバイスである。表示デバイス 2 0 7 は、タッチスクリーンを含んでもよい。G P U (Graphics Processing Unit) 2 0 8 は、同種の多数の演算を高速且つ並列的に実行可能な、学習処理に専用のプロセッサである。G P U 2 0 8 は、C P U 2 0 1 と連携して動作し、例えば後述する提案部 2 3 0 及び学習部 2 5 0 の機能のための演算の一部を実行する。

#### 【 0 0 2 0 】

##### < 1 - 3 . 課題 >

30

ユーザ端末 1 0 0 のような端末装置は、近年ますます多機能化及び高機能化している。端末装置のアプリケーションは、年齢、言語、嗜好その他の属性の異なる様々なユーザにとってユーザビリティが最適となるように、いくつかのユーザ支援機能を提供し得る。例えば、ウェブブラウザ及びメーラのようなテキストを表示するアプリケーションは、ユーザによるテキストの内容の把握を支援するための、辞書機能及び言語設定機能を提供するかもしれない。画像又は動画を再生するコンテンツプレーヤは、ユーザによるコンテンツの閲覧を支援するための、輝度変更機能を提供するかもしれない。しかし、端末装置の機能を自在に使いこなすことのできるユーザは少なく、機能の存在自体にユーザが気付かないケースもある。

#### 【 0 0 2 1 】

40

ここで、システムがユーザの顔画像を分析してユーザの表情の変化を認識し、認識した変化に基づいてユーザに適していると推定される機能をユーザへ提案する、というアプローチが考えられる。しかし、ユーザの表情の変化には個人差がある。そのため、単に既知のモデル又はパターンに従って顔画像を分析するだけでは、表情認識の十分な精度を確保できず、ユーザに適した機能を適切に提案することが難しい。そこで、本実施形態に係る情報処理システム 1 0 は、次のような仕組みによって、表情認識に基づく機能提案の手法を改善する。

#### 【 0 0 2 2 】

##### < 1 - 4 . 基本的な仕組み >

図 3 は、情報処理システム 1 0 において実行され得る提案処理の概略的な流れの一例を

50

示すシーケンス図である。図 3 に示した提案処理には、ユーザ端末 1 0 0 及び提案サーバ 2 0 0 が関与する。

【 0 0 2 3 】

まず、S 3 0 1 で、ユーザ端末 1 0 0 は、ユーザ端末 1 0 0 にインストールされているアプリケーション ( A P ) のうちの 1 つを実行する。実行されたアプリケーションは、表示デバイス 1 0 7 にアプリケーション画面を表示させる。ユーザは、アプリケーション画面上のコンテンツを閲覧し、又はアプリケーション画面を操作する。ユーザは、閲覧又は操作に不都合又は不満を感じた場合には、表情を変化させるものと予期される。

【 0 0 2 4 】

S 3 0 2 で、ユーザ端末 1 0 0 は、カメラ 1 0 5 を用いてユーザを撮像し、撮像画像を生成する。ここでは、ユーザの顔を写した顔画像が生成されるものとする。ユーザ端末 1 0 0 は、顔画像を分析することにより、ユーザの表情の変化の度合いを定量的に評価する。その評価の結果算出される評価値がある条件を満たす場合、ユーザ端末 1 0 0 は、S 3 0 3 で、提案要求を提案サーバ 2 0 0 へ送信する。提案要求は、カメラ 1 0 5 により撮像されたユーザの顔画像データを含み得る。提案要求は、提案サーバ 2 0 0 により受信される。

【 0 0 2 5 】

S 3 0 4 で、提案サーバ 2 0 0 は、受信した提案要求に含まれるユーザの顔画像データを学習済みモデルに適用することにより、ユーザ端末 1 0 0 において起動すべき機能を決定する。例えば、ウェブブラウザを閲覧している最中に未知の単語に遭遇したユーザは、眉をひそめる表情をするかもしれない。その場合、提案サーバ 2 0 0 は、単語の意味を検索して表示する辞書機能を起動することを提案し得る。S 3 0 5 で、提案サーバ 2 0 0 は、決定した提案機能を特定する提案応答をユーザ端末 1 0 0 へ送信する。提案応答は、ユーザ端末 1 0 0 により受信される。

【 0 0 2 6 】

S 3 0 6 で、ユーザ端末 1 0 0 は、受信した提案応答により特定される提案機能を起動する。ユーザは、提案機能を起動した結果、自らの不都合又は不満が解消した場合には、表情をもとに戻す ( 又はより好意的な表情に変化させる ) ものと予期される。

【 0 0 2 7 】

S 3 0 7 で、ユーザ端末 1 0 0 は、カメラ 1 0 5 を用いて再びユーザを撮像し、ユーザの顔画像を生成する。ユーザ端末 1 0 0 は、顔画像を分析することにより、ユーザの表情の変化の度合いを定量的に評価する。そして、提案機能の起動の前後のユーザの表情の変化度がある条件を満たす場合、ユーザ端末 1 0 0 は、S 3 0 8 で、学習済みモデルの再学習を求める学習要求を提案サーバ 2 0 0 へ送信する。学習要求は、提案サーバ 2 0 0 により受信される。

【 0 0 2 8 】

S 3 0 9 で、提案サーバ 2 0 0 は、学習要求の受信に応じて、S 3 0 4 で使用した学習済みモデルの再学習を実行する。例えば、提案サーバ 2 0 0 は、機能を提案した結果としてユーザの不都合又は不満が解消した場合には、提案した機能を正解とし、S 3 0 3 で受信した顔画像データにより示される顔画像を学習用画像として用いて、学習済みモデルの再学習を実行する。

【 0 0 2 9 】

このように、ユーザの実際の顔画像を用いた再学習を反復的に実行することで、学習済みモデルがユーザの表情の変化を適応的に学習し、よりユーザに適した機能を提案サーバ 2 0 0 が提案することが可能となる。こうした再学習の仕組みを実現するためのユーザ端末 1 0 0 及び提案サーバ 2 0 0 の構成について、次節以降でより具体的に説明する。

【 0 0 3 0 】

< < 2 . 各装置の具体的な構成 > >

< 2 - 1 . ユーザ端末 >

図 4 ( A ) は、本実施形態に係るユーザ端末 1 0 0 の機能面の構成の一例を示すブロッ

10

20

30

40

50



ク図である。図 4 ( A ) を参照すると、ユーザ端末 1 0 0 は、画像取得部 1 2 0、表情評価部 1 3 0、記憶部 1 4 0、N 個のアプリケーション部 1 5 0 - 1 ~ 1 5 0 - N 及び実行制御部 1 6 0 を備える。なお、ユーザ端末 1 0 0 は、図 4 ( A ) に示したものの以外の多くの機能ブロックをさらに有し得るが、説明を曖昧にしないために、ここでは機能の提案に関係しない機能ブロックについての図示及び説明を省略する。

#### 【 0 0 3 1 】

画像取得部 1 2 0 は、カメラ 1 0 5 により撮像されるユーザの顔画像を取得する。画像取得部 1 2 0 は、周期的にユーザの顔画像を取得してもよく、又は入力デバイス 1 0 6 への所定のユーザ入力若しくはマイクロフォン 1 0 8 への所定の音声入力が検知された場合に、ユーザの顔画像を取得してもよい。画像取得部 1 2 0 は、取得した顔画像を表す顔画像データを表情評価部 1 3 0 及び実行制御部 1 6 0 へ出力する。

10

#### 【 0 0 3 2 】

表情評価部 1 3 0 は、ユーザの表情の変化を定量的に認識するための表情評価値を、画像取得部 1 2 0 により取得されるユーザの顔画像に基づいて算出する。本実施形態において、表情評価部 1 3 0 は、記憶部 1 4 0 に予め記憶される評価モデル 1 4 2 にユーザの顔画像を適用することにより、表情評価値を取得する。一例として、評価モデル 1 4 2 は、機械学習を通じて予め獲得された学習済みモデルであってもよい。

#### 【 0 0 3 3 】

図 5 ( A ) は、評価モデル 1 4 2 の入出力データの例を説明するための説明図である。ここでは、評価モデル 1 4 2 は、ニューラルネットワークモデルであるものとする。図 5 ( A ) を参照すると、評価モデル 1 4 2 は、入力層、複数の中間層及び出力層を含む。評価モデル 1 4 2 は、ユーザの顔画像データ ( 2 次元的に配列された画素の画素値データ ) を入力データとして受け取る。そして、評価モデル 1 4 2 は、例えば [ 0 , 1 ] というレンジ内の値である表情評価値を、ニューラルネットワークモデルの演算の結果として出力する。

20

#### 【 0 0 3 4 】

学習段階では、評価モデル 1 4 2 は、複数の人間の顔画像 ( サンプル画像 ) を学習用画像として受け付ける。各学習用画像には、当該画像に映るユーザが表情に示している満足度が、教師データとして予めラベル付けされる。満足度もまた、[ 0 , 1 ] というレンジ内の値である。満足度は、各学習用画像を見たエンジニアにより主観的にラベル付けされてもよい。その代わりに、例えば顔領域と特徴点 ( 例えば、目、鼻、口及び眉毛の頂点 ) との間の位置関係、目尻及び口元が描く曲線の曲率、並びに目領域に占める黒目の割合といった画像特徴量に基づいて、満足度が採点されてもよい。概して、アプリケーションの状況に不都合を感じておらず満足しているユーザの表情 ( 例えば、笑顔又はうなずいた顔 ) には、より高い評価値 ( 例えば、1 に近い値 ) が付与され得る。逆に、不都合又は不満を感じているユーザの表情 ( 例えば、しかめっ面又は俯いた顔 ) には、より低い評価値 ( 例えば、ゼロに近い値 ) が付与され得る。学習段階において、評価モデル 1 4 2 への学習用画像の入力、モデルパラメータに従った表情評価値の算出、及び教師データに対する評価値の誤差に基づくモデルパラメータの修正が、多数の学習用画像について反復される。十分な回数の反復の結果、評価モデル 1 4 2 は、ユーザの顔画像に基づいて表情の変化を定量的に表す表情評価値を算出するための、ある程度正確なモデルとなる。なお、典型的には、学習段階の処理は、ユーザ端末 1 0 0 とは別個の装置によって行われる。そして、ユーザ端末 1 0 0 は、既に獲得された学習済みモデルを記憶部 1 4 0 にダウンロードして利用し得る。但し、当然ながら、ユーザ端末 1 0 0 において学習段階の処理が実行されてもよい。

30

40

#### 【 0 0 3 5 】

なお、評価モデル 1 4 2 は、図 5 ( A ) を用いて説明したような機械学習を通じて学習された学習済みモデルには限定されない。例えば、表情評価部 1 3 0 は、学習済みモデルの代わりに、画像特徴量と表情評価値との間の数学的関係を記述した数式モデルなど、他の種類のモデルを用いて表情評価値を算出してもよい。

50

## 【 0 0 3 6 】

本実施形態において、表情評価部 1 3 0 は、提案サーバ 2 0 0 への提案要求の送信の前（即ち、提案機能の起動の前。図 3 の S 3 0 2 ）に第 1 の表情評価値を取得し、及び、提案機能の起動の後（図 3 の S 3 0 7 ）に第 2 の表情評価値を取得する。第 1 の表情評価値は、提案要求の送信の前に撮像されたユーザの第 1 の顔画像を評価モデル 1 4 2 に適用することにより算出される。第 2 の表情評価値は、提案機能の起動の後に撮像されたユーザの第 2 の顔画像を評価モデル 1 4 2 に適用することにより算出される。第 1 の表情評価値から第 2 の表情評価値への変化は、提案機能の起動によって変化したユーザの表情の変化度に相当し得る。後述する実行制御部 1 6 0 は、このユーザの表情の変化度に依存して、学習要求を提案サーバ 2 0 0 へ送信する。

10

## 【 0 0 3 7 】

記憶部 1 4 0 は、画像取得部 1 2 0 により取得されるユーザの顔画像データ、上述した評価モデル 1 4 2 、ユーザを識別するユーザ識別情報、及び稼働中のアプリケーションを識別する A P 識別情報を記憶する。また、記憶部 1 4 0 は、後に表 1 ～表 3 に例示するアプリケーションリスト、機能リスト及びマッピング情報を記憶する。

## 【 0 0 3 8 】

アプリケーション部 1 5 0 - 1 ～ 1 5 0 - N は、ユーザ端末 1 0 0 の様々なアプリケーションにそれぞれ対応する機能ブロックである。例えば、アプリケーション部 1 5 0 - 1 はウェブブラウザ、アプリケーション部 1 5 0 - 2 はメーラ、アプリケーション部 1 5 0 - 3 はインスタントメッセンジャ、アプリケーション部 1 5 0 - 4 はコンテンツプレーヤに相当し得る。但し、これらは単なる例示に過ぎず、ユーザ端末 1 0 0 はいかなるアプリケーションの組合せを有していてもよい。

20

## 【 0 0 3 9 】

ユーザ端末 1 0 0 により実行可能なアプリケーションの各々は、A P 識別情報（例えば、A P I D ）により一意に識別される。また、ユーザ端末 1 0 0 において起動可能な機能は、機能識別情報（例えば、機能 I D ）により一意に識別される。次の表 1 は、A P I D により識別されるアプリケーションのリストの一例を示している。

## 【 0 0 4 0 】

【表 1】

AP ID	アプリケーション名
A1	ウェブブラウザ
A2	メーラ
A3	インスタントメッセンジャ
A4	フォトアルバム
A5	動画プレーヤ

表 1. アプリケーションリストの一例

30

## 【 0 0 4 1 】

次の表 2 は、機能 I D により識別される機能のリストの一例を示している。表 1 及び表 2 から理解されるように、あるアプリケーション（例えば、ウェブブラウザ）が他のアプリケーションの稼働中に起動可能な機能の 1 つとして位置付けられてもよい。

40

## 【 0 0 4 2 】

50

【表 2】

機能ID	機能名
F1	辞書
F2	言語設定
F3	輝度変更
F4	カメラ
F5	ウェブブラウザ

表 2. 機能リストの一例

10

## 【0043】

ユーザ端末100及び提案サーバ200は、各アプリケーションの稼働中にどの機能が起動可能であるかを示すマッピング情報を予め共有する。また、一実施例において、特定の種類のアプリケーションについては、当該アプリケーションの稼働中のシステムからの提案に基づく機能の起動は無効化されてもよい。どの種類のアプリケーションについて機能の提案が無効化されるかは、予め固定的に定義されてもよく、又はユーザにより設定可能であってもよい。次の表3は、各アプリケーションと当該アプリケーションの稼働中に起動可能な機能とを関連付けるマッピング情報の一例を示している。

## 【0044】

【表 3】

20

AP ID	提案	各機能の起動可否				
		F1	F2	F3	F4	F5
A1	有効	○	○	○		
A2	有効	○	○	○	○	
A3	有効	○	○	○	○	○
A4	有効			○	○	
A5	無効					

表 3. マッピング情報の一例

30

## 【0045】

表3の例では、5種類のアプリケーションのうち、AP ID = “A5”である動画プレーヤ以外の4種類のアプリケーションについて、システムからの提案に基づく機能の起動が有効化されている。ウェブブラウザ(A1)については、辞書機能(F1)、言語設定機能(F2)及び輝度変更機能(F3)が起動を提案される機能の候補である。メーラ(A2)については、辞書機能(F1)、言語設定機能(F2)、輝度変更機能(F3)及びカメラ機能(F4)が起動を提案される機能の候補である。インスタントメッセージング(A3)については、辞書機能(F1)、言語設定機能(F2)、輝度変更機能(F3)、カメラ機能(F4)及びウェブブラウザ機能(F5)が起動を提案される機能の候補である。フォトアルバム(A4)については、輝度変更機能(F3)及びカメラ機能(F4)が起動を提案される機能の候補である。

40

## 【0046】

実行制御部160は、ユーザ端末100におけるアプリケーションの実行及び関連する機能の実行を制御する。例えば、実行制御部160は、入力デバイス106を介してユーザにより指定されたアプリケーションを実行する。また、実行制御部160は、アプリケーションの稼働中に表情評価部130により取得される表情評価値を監視する。例えば、実行制御部160は、表情評価値が提案要求のための閾値(以下、提案要求閾値という)を下回る場合、ユーザ端末100において起動すべき機能の提案を求める提案要求を、通信I/F111を介して提案サーバ200へ送信する。実行制御部160は、提案サーバ200へ送信する提案要求に、ユーザの顔を映した顔画像データを含める。そして、実行

50

制御部 160 は、提案要求に応じて提案サーバ 200 により決定される提案機能を特定する提案応答を、通信 I/F 111 を介して提案サーバ 200 から受信する。提案機能は、提案サーバ 200 において、提案要求に含まれる顔画像データを、後述する学習済みモデル 242 に適用することにより決定される。

#### 【0047】

本実施形態において、実行制御部 160 は、ユーザ端末 100 において稼働中のアプリケーションを識別する AP ID を提案要求にさらに含める。これは、提案サーバ 200 により決定される提案機能を、稼働中のアプリケーションに依存して異なるものとするためである。

#### 【0048】

図 5 (B) は、提案サーバ 200 により保持される学習済みモデル 242 の入出力データの例を説明するための説明図である。ここでは、学習済みモデル 242 は、ニューラルネットワークモデルであるものとする。図 5 (B) を参照すると、学習済みモデル 242 は、入力層、複数の中間層及び出力層を含む。学習済みモデル 242 は、ユーザの顔画像データ及び AP ID を入力データとして受け取る。ここでの AP ID は、ユーザ端末 100 において稼働中のアプリケーションを識別する識別情報である。学習済みモデル 242 は、提案機能ベクトルを、ニューラルネットワークモデルの演算の結果として出力する。例えば、表 2 の機能リストには 5 個の機能が列挙されている。この場合、提案機能ベクトルは、5 次元のベクトルであり得る。提案機能ベクトルの 5 個の要素値は、例えば出力層におけるソフトマックス関数の演算の結果として、[ 0 , 1 ] のレンジ内の値を示し、合計が 1 となるように正規化される。各要素の値は、対応する機能を起動することがユーザに適している度合いを示す。例えば、提案機能ベクトルの 1 番目の要素が最も高い値を示す場合、機能リストの 1 番目の辞書機能 (F1) が提案機能として決定され得る。提案機能ベクトルの 2 番目の要素が最も高い値を示す場合、機能リストの 2 番目の言語設定機能 (F2) が提案機能として決定され得る。その他の要素の値の意味も同様である。なお、表 3 に示したマッピング情報において、提案要求により指定された AP ID に関連付けられていない (起動可能ではない) 機能は、提案機能の決定において除外され得る。

#### 【0049】

一変形例において、提案機能ベクトルは、「提案なし」という結論のための 1 つの要素を追加的に含んでもよい。このケースでは、機能リストに 5 個の機能が列挙されていれば、学習済みモデル 242 により出力される提案機能ベクトルは、6 次元のベクトルとなる。そして、提案機能ベクトルにおいて「提案なし」に対応する要素が最も高い値を示すときは、提案サーバ 200 は、どの機能も提案機能としないことを決定し得る。

#### 【0050】

他の変形例において、学習済みモデル 242 は、アプリケーションごとに別々に提供されてもよい。この場合、あるアプリケーションのための学習済みモデル 242 により出力される提案機能ベクトルは、表 3 のマッピング機能において当該アプリケーションに関連付けられている機能の数に対応する次元を有し得る。

#### 【0051】

学習段階では、学習済みモデル 242 は、複数の人間の顔画像 (サンプル画像) と AP ID との組合せを学習用データとして受け付ける。各サンプル画像は、サンプルユーザが所望の機能を起動した時点の直前に撮像された顔画像であり、各 AP ID はその時点で稼働中であったアプリケーションを識別する。これら顔画像及び AP ID のペアに、サンプルユーザにより起動された機能について「1」、他の候補機能について「0」を示す提案機能ベクトルが教師データとして付与される。学習段階において、学習済みモデル 242 への学習用データの入力、モデルパラメータに従った提案機能ベクトルの算出、及び教師データに対する提案機能ベクトルの誤差に基づくモデルパラメータの修正が、多数の学習用データについて反復される。十分な回数の反復の結果、学習済みモデル 242 は、ユーザの顔画像及び AP ID に基づいて提案機能を決定するためのモデルとなる。

#### 【0052】

10

20

30

40

50

本実施形態において、学習済みモデル 242 は、提案サーバ 200 によりユーザごとに保持され得る。学習済みモデル 242 は、再学習が一度も実行されていない導入の当初においては、全てのユーザに共通のモデルパラメータ値のセットで構成され得る。しかし、後述する再学習が繰返される結果として、学習済みモデル 242 は、個々のユーザの表情の個人差を反映したものとなる。こうしたユーザ固有の学習済みモデル 242 へのアクセスを可能とするために、実行制御部 160 は、ユーザを識別するユーザ識別情報（例えば、ユーザ ID）をも提案要求に含めて提案サーバ 200 へ送信し得る。

#### 【0053】

実行制御部 160 は、提案サーバ 200 からの提案応答の受信に応じて、提案応答により特定された提案機能を起動する。実行制御部 160 は、提案機能を起動する前に、提案機能の起動の承認をユーザに求めてもよい。図 6 は、提案機能の起動の承認をユーザに求めるためのユーザインタフェースの一例を示す説明図である。図 6 を参照すると、ユーザ端末 100 においてアプリケーション画面 600 が表示されており、アプリケーション画面 600 の右下にポップアップウィンドウ 610 が重畳されている。ポップアップウィンドウ 610 は、提案機能として特定された辞書機能を起動することについてユーザに承認を求めるメッセージを示している。ユーザは、例えばポップアップウィンドウ 610 のボタン 611 を選択することで辞書機能の起動を承認し、又はボタン 612 を選択することで辞書機能の起動を拒否することができる。実行制御部 160 は、このようなインタラクションの結果として、提案機能の起動がユーザにより承認された場合に、提案機能を起動し得る。

#### 【0054】

なお、実行制御部 160 は、図 6 に示したようなポップアップウィンドウではなく、他のユーザインタフェースを通じてユーザに提案機能の起動の承認を求め又は提案機能の起動を促してもよい。例えば、実行制御部 160 は、提案機能の起動の指示を受け付けるためのアイコン若しくはショートカットリンクを表示デバイス 107 に表示させてもよく、又は提案機能を紹介する音声メッセージをスピーカ 109 に再生させてもよい。

#### 【0055】

実行制御部 160 は、提案機能を起動した後、引続き表情評価部 130 により取得される表情評価値を監視する。そして、実行制御部 160 は、ユーザの表情の変化度が所定の条件を満たす場合に、ユーザの顔画像データを用いた学習済みモデル 242 の再学習を求める学習要求を、通信 I/F 111 を介して提案サーバ 200 へ送信する。ここでの所定の条件とは、例えば、提案機能の起動の前に取得された第 1 の表情評価値に対する提案機能の起動の後に取得された第 2 の表情評価値の差が学習要求のための閾値（以下、学習要求閾値という）を上回ることを含み得る。ここでの第 1 の表情評価値は、提案要求の送信をトリガした、提案要求閾値よりも低い評価値であり得る。第 2 の表情評価値が第 1 の表情評価値から学習要求閾値を超えて上昇したという変化は、提案機能の起動によってユーザの満足度が有意に改善したことを意味し得る。こうしたケースでは、起動した提案機能が、ユーザに実際に適したものであったとの推定が成り立つ。そのため、提案要求に含まれていた顔画像データにより示される顔画像と AP ID とのペアを学習用データとして用いて再学習を実行することで、次回以降の処理で学習済みモデル 242 に基づいて適切な機能の起動が提案される可能性を高めることができる。この再学習においては、（表情評価値を改善させた）提案機能について「1」、他の候補機能について「0」を示す提案機能ベクトルが教師データとして利用され得る。実行制御部 160 は、学習要求にもユーザ ID を含める。学習済みモデルの再学習は、提案サーバ 200 において、このユーザ ID に基づいて、ユーザ固有のモデルを獲得するためにユーザごとに実行され得る。

#### 【0056】

実行制御部 160 は、提案機能を起動した後、ユーザの表情の変化度が上記所定の条件を満たさない場合には、例えばユーザの満足度が有意に改善しなかったことを示す提案結果通知を提案サーバ 200 へ送信し得る。実行制御部 160 は、提案機能の起動が拒否された場合には、提案機能の起動が拒否されたことを示す提案結果通知を提案サーバ 200

10

20

30

40

50

へ送信し得る。これら提案結果通知は、提案機能がユーザにとって適したものではなかったことを提案サーバ200へ示唆する。そこで、提案サーバ200は、提案結果通知に基づいて学習済みモデルの再学習を実行してもよい。この場合の再学習の教師データは、例えば全ての候補機能について「0」を、「提案なし」について「1」を示す提案機能ベクトルであってもよい。代替的に、提案サーバ200は、提案結果通知に基づいて学習済みモデルの再学習を実行しなくてもよい。

#### 【0057】

実行制御部160は、ユーザ端末100において稼働中のアプリケーションが提案に基づく機能の起動を無効化されたアプリケーションである場合に、提案サーバ200に提案要求を送信しなくてもよく、又は提案された機能を起動しなくてもよい。例えば、ユーザが動画プレーヤを実行して動画コンテンツを視聴中である場合、ポップアップウィンドウを表示してその視聴を妨げることは、ユーザにとって好ましくないと想定される。ユーザがゲームアプリケーションを実行してゲームを楽しんでいる最中についても同様である。したがって、そうした状況においてシステムからの提案を差し控えることで、ユーザにとっての利便性が却って低下することを回避することができる。

#### 【0058】

実行制御部160は、システムからの機能の提案の有効化及び無効化を、アプリケーションの種別ごとではなくコンテンツごとに制御してもよい。例えば、実行制御部160は、動画プレーヤの稼働中において、特定の動画コンテンツが再生されているときに、提案サーバ200への提案要求の送信を差し控えてもよい。どの種類のコンテンツについて機能の提案を有効化し又は無効化するかが、ユーザにより設定されてもよい。例えば、ユーザが深く没入するようなコンテンツが再生されている状況では、ユーザの表情はコンテンツの内容に強く影響され、ユーザビリティとは関係なく大きく変化する可能性が高い。そのため、そうした状況において表情の変化に基づく機能の提案を無効化することで、表情の無用な評価のためのリソースの浪費や不適切な機能の提案といった望ましくない事態を回避することができる。

#### 【0059】

##### < 2 - 2 . 提案サーバ >

図4(B)は、本実施形態に係る提案サーバ200の機能面の構成の一例を示すブロック図である。図4(B)を参照すると、提案サーバ200は、要求受付部220、提案部230、記憶部240及び学習部250を備える。

#### 【0060】

要求受付部220は、ユーザ端末100において起動すべき機能の提案を求める提案要求を、通信I/F205を介してユーザ端末100から受信する。提案要求は、上述したように、ユーザ端末100のユーザの顔画像を示す顔画像データを含む。提案要求は、さらに、ユーザ端末100において稼働中のアプリケーションを識別するAP ID、及びユーザを識別するユーザIDを含み得る。要求受付部220は、受け付けた提案要求を提案部230へ出力する。

#### 【0061】

提案部230は、要求受付部220により受け付けられた提案要求に含まれる顔画像データを学習済みモデル242に適用することにより、ユーザ端末100において起動すべき提案機能を決定する。ここでの学習済みモデル242は、ユーザ端末100の導入の当初においては複数のユーザに共通の既定のモデルであり、但し上述した再学習の繰返しを通じてユーザ固有のモデルとなる。そのため、提案部230は、提案要求に含まれるユーザIDに関連付けられている学習済みモデル242を、提案機能を決定する際に利用する。提案要求に含まれるAP IDもまた、学習済みモデル242への入力パラメータであり得る。結果的に、提案部230により決定される提案機能は、AP IDにより識別される稼働中のアプリケーションに依存して異なる。提案部230は、決定した提案機能を特定する(例えば、提案機能を識別する機能IDを含む)提案応答を要求受付部220へ出力する。提案応答は、要求受付部220から通信I/F205を介してユーザ端末100

10

20

30

40

50

0へ送信される。このようにして、提案部230は、ユーザ端末100に提案機能を起動させる。

#### 【0062】

記憶部240は、ユーザの顔画像データに基づいてユーザ端末100において起動すべき提案機能を決定するための学習済みモデル242を記憶する。例えば、記憶部240は、複数のユーザに共通の既定のモデルパラメータのセットと、後述する学習部250による再学習を通じて獲得された各ユーザに固有のモデルパラメータのセットとを含み得る。

#### 【0063】

要求受付部220は、提案応答がユーザ端末100へ送信された後、ユーザ端末100からの上述した学習要求又は提案結果通知の受信を待ち受ける。要求受付部220は、学習要求が通信I/F205を介して受信された場合、学習要求を学習部250へ出力する。また、要求受付部220は、提案結果通知が受信された場合、提案結果通知を学習部250へ出力する。

#### 【0064】

学習部250は、学習済みモデルの再学習を求める学習要求がユーザ端末100から受信された場合に、学習済みモデル242の再学習を実行する。学習部250は、その再学習において、学習要求に先立ってユーザ端末100へ提案した機能を正解とし、提案機能の決定に際して学習済みモデル242に適用した顔画像を学習用画像として用いる。学習用データは、この学習用画像と提案要求に含まれていたAPIDとのペアである。教師データの提案機能ベクトルは、提案機能について「1」、他の候補機能について「0」を示し得る。学習部250は、学習要求に含まれ得るユーザIDに基づいて、再学習をユーザごとに実行する。即ち、学習部250は、指定されたユーザIDに関連付けて記憶部240に記憶されている学習済みモデル242のモデルパラメータのセットを再学習を通じて更新し得る。こうした再学習の反復を通じて、提案サーバ200が学習済みモデル242に基づいて個々のユーザに実際に適した提案機能を適切に決定する能力が継続的に高められる。学習部250は、提案結果通知がユーザ端末100から受信された場合には、上述したように、学習済みモデル242の再学習を実行してもしなくてもよい。

#### 【0065】

<<3. 処理の流れ>>

<3-1. ユーザ端末の処理>

図7は、本実施形態に係るユーザ端末100により実行される処理の流れの一例を示すフローチャートである。図7に示した処理は、例えば、ROM102からRAM103へロードされるコンピュータプログラムをコントローラ101が実行することにより実現され得る。なお、以下の説明では、処理ステップをS(ステップ)と略記する。

#### 【0066】

フローチャートの開始時に、ユーザ端末100において、何らかのアプリケーションが稼働しているものとする。S701で、実行制御部160は、ユーザへ機能を提案するための提案モードが有効であるか否かを判定する。例えば、実行制御部160は、稼働しているアプリケーションが提案モードが無効化されたアプリケーションではない場合に、提案モードが有効であると判定してもよい。その代わりに、実行制御部160は、再生中のコンテンツが提案モードが無効化されたコンテンツではない場合に、提案モードが有効であると判定してもよい。提案モードが有効であると判定された場合、処理はS703へ進む。一方、提案モードが無効であると判定された場合、図7のフローチャートは終了する。

#### 【0067】

S703で、実行制御部160は、ユーザによる操作を監視する。例えば、入力デバイス106への所定のユーザ入力が発知された場合、処理はS705へ進み、実行制御部160から表情評価部130へユーザの表情の評価が指示される。なお、上述したように、ユーザの表情の評価は、ユーザ入力に関わらず周期的に行われてもよい。S705で、画像取得部120は、カメラ105により撮像されるユーザの顔画像を取得する。次いで、S707で、表情評価部130は、画像取得部120により取得されたユーザの顔画像を

10

20

30

40

50

評価モデル 142 に適用することにより、表情評価値（第 1 の表情評価値）を取得する。次いで、S709 で、実行制御部 160 は、表情評価部 130 により取得された表情評価値が提案要求閾値を下回るか否かを判定する。表情評価値が提案要求閾値を下回ると判定された場合、処理は S711 へ進む。一方、表情評価値が提案要求閾値を下回らないと判定された場合、処理は S703 へ戻る。

【0068】

実行制御部 160 は、表情評価値が提案要求閾値を下回ると判定した場合、S711 で、ユーザ端末 100 において起動すべき機能の提案を求める提案要求を、通信 I/F 111 を介して提案サーバ 200 へ送信する。提案要求は、画像取得部 120 により取得された顔画像データ、稼働しているアプリケーションを識別する AP ID、及びユーザを識別するユーザ ID を含む。次いで、S713 で、実行制御部 160 は、提案サーバ 200 からの提案応答の受信を待ち受ける。提案サーバ 200 からの提案応答が受信されると、処理は S715 へ進む。

10

【0069】

S715 で、実行制御部 160 は、提案サーバ 200 から受信された提案応答により特定された提案機能の起動の承認を、例えば図 6 を用いて説明したポップアップウィンドウ 610 を表示することにより、ユーザに求める。その後の処理は、提案機能の起動がユーザにより承認されたか否かに依存して分岐する。提案機能の起動が承認された場合、処理は S719 へ進む。一方、提案機能の起動が承認されなかった場合、処理は S729 へ進む。

20

【0070】

提案機能の起動が承認された場合、S719 で、実行制御部 160 は、提案応答により特定された提案機能を起動する。次いで、S721 で、画像取得部 120 は、カメラ 105 により撮像されるユーザの顔画像を取得する。次いで、S723 で、表情評価部 130 は、画像取得部 120 により取得されたユーザの顔画像を評価モデル 142 に適用することにより、表情評価値（第 2 の表情評価値）を取得する。

【0071】

次いで、S725 で、実行制御部 160 は、表情評価部 130 により取得された表情評価値の変化（例えば、第 1 の表情評価値に対する第 2 の表情評価値の差）が学習要求閾値を上回るかを判定する。表情評価値の変化が学習要求閾値を上回ると判定された場合、処理は S727 へ進む。一方、表情評価値の変化が学習要求閾値を上回らないと判定された場合、処理は S729 へ進む。

30

【0072】

S727 で、実行制御部 160 は、ユーザの表情の変化度が学習要求閾値を上回るという条件が満たされたため、学習済みモデル 242 の再学習を求める学習要求を、通信 I/F 111 を介して提案サーバ 200 へ送信する。一方、S729 で、実行制御部 160 は、提案機能の起動が拒否されたこと、又は提案機能がユーザにとって適したものではなかったことを通知するための提案結果通知を、通信 I/F 111 を介して提案サーバ 200 へ送信する。その後、図 7 のフローチャートは終了する。

【0073】

< 3 - 2 . 提案サーバの処理 >

図 8 は、本実施形態に係る提案サーバ 200 により実行される処理の流れの一例を示すフローチャートである。図 8 に示した処理は、例えば、HDD 204 から RAM 203 へロードされるコンピュータプログラムを CPU 201 が実行することにより実現され得る。

【0074】

まず、S801 で、要求受付部 220 は、ユーザ端末 100 において起動すべき機能の提案を求める提案要求のユーザ端末 100 からの受信を待ち受ける。通信 I/F 205 を介して提案要求が受信されると、処理は S803 へ進む。提案要求は、上述したように、ユーザの顔画像データ、ユーザ端末 100 において稼働しているアプリケーションを識別する AP ID、及びユーザを識別するユーザ ID を含み得る。

40

50



## 【 0 0 7 5 】

S 8 0 3 で、提案部 2 3 0 は、提案要求に含まれていた顔画像データ（及び A P I D）を、（ユーザ I D に関連付けて記憶されている）学習済みモデル 2 4 2 に適用することにより、ユーザ端末 1 0 0 において起動すべき提案機能を決定する。次いで、S 8 0 5 で、提案部 2 3 0 は、決定した提案機能を特定する提案応答を、要求受付部 2 2 0 からユーザ端末 1 0 0 へ送信させる。

## 【 0 0 7 6 】

次いで、S 8 0 7 及び S 8 0 9 で、要求受付部 2 2 0 は、ユーザ端末 1 0 0 からの学習要求又は提案結果通知の受信を待ち受ける。学習要求が受信された場合、S 8 1 1 で、学習部 2 5 0 は、S 8 0 3 で決定した提案機能を正解とし、提案機能の決定に際して使用した顔画像データ（及び A P I D）を学習用データとして用いて、ユーザ固有の学習済みモデル 2 4 2 の再学習を実行する。図 8 の例では、学習部 2 5 0 は、提案結果通知が受信された場合、再学習をスキップする。その後、図 8 のフローチャートは終了する。

## 【 0 0 7 7 】

なお、学習部 2 5 0 は、S 8 1 1 で、提案機能を正解とする教師データに関連付けて学習用データを記憶部 2 4 0 に一時的に記憶させ、ある程度の量の学習用データが蓄積されたタイミングでまとめて再学習を実行してもよい。また、学習部 2 5 0 は、提案結果通知が受信された場合にも、再学習を実行してもよい。

## 【 0 0 7 8 】

< < 4 . ま と め > >

ここまで、図 1 ~ 図 8 を用いて、本開示の実施形態について詳細に説明した。上述した実施形態では、端末装置において起動すべき機能が、ユーザの顔画像データを含む提案要求に応じて、サーバ装置において当該顔画像データを学習済みモデルに適用することにより決定される。そして、上記サーバ装置により決定された提案機能が、上記端末装置により起動される。上記端末装置は、上記提案機能の起動の前後のユーザの表情の変化度が所定の条件を満たす場合に、上記顔画像データを用いた上記学習済みモデルの再学習を上記サーバ装置へ要求する。かかる構成によれば、提案機能の決定のための学習済みモデルを、実用の場面における実際のユーザの表情に基づく再学習を通じて強化することができる。それにより、ユーザの表情に基づいてより適切な機能をユーザに提案することが可能となる。また、ユーザが比較的少ない計算リソースしか有しないユーザ端末を使用している場合にも、上記サーバ装置からの機能の提案によってユーザを支援することが可能である。

## 【 0 0 7 9 】

また、上述した実施形態では、上記学習要求は、上記ユーザを識別するユーザ識別情報を含み、上記学習済みモデルの上記再学習は、ユーザ固有のモデルを獲得するためにユーザごとに実行され得る。かかる構成によれば、ユーザの表情の変化の個人差が再学習を通じて個々の学習済みモデルに反映されるため、再学習後に提案される機能がより確実に各ユーザにふさわしいものとなる。

## 【 0 0 8 0 】

また、上述した実施形態では、表情の変化の評価値を算出するための評価モデルが提供される。そして、ユーザの表情の上記変化度は、上記提案機能の起動の前後で取得される 2 つの顔画像を上記評価モデルに適用することにより取得される 2 つの評価値の間の変化に相当し得る。かかる構成によれば、表情の変化の度合いの定量的な評価に基づいて、学習済みモデルの効果的な再学習を進めることができる。例えば、2 つの評価値の間の差が第 1 の閾値を上回る場合、又は、2 つの評価値の間の変化がユーザの満足度が改善したことを示す場合に、学習済みモデルの再学習が実行され得る。これらの場合には、上記提案機能がユーザに実際に適したものであったとの推定が成り立つことから、上記提案機能を提案すべき機能の正解として用いて再学習を実行することで、次回以降も適切な機能が提案される可能性を高めることができる。

## 【 0 0 8 1 】

また、上述した実施形態では、上記評価モデルにユーザの顔画像を適用することにより

10

20

30

40

50

取得される評価値が第2の閾値を下回る場合に、上記提案要求が上記サーバ装置へ送信され得る。かかる構成によれば、ユーザが例えばアプリケーションの操作又はコンテンツの閲覧に不都合又は不満を感じている状況を顔画像に現れるユーザの表情から把握し、ユーザに適した機能の起動を適時に提案することができる。また、そうした状況における機能の提案の結果（例えば、ユーザの不都合又は不満が解消された、など）に基づいて学習済みモデルを効果的に強化することができる。

#### 【0082】

また、上述した実施形態では、上記提案要求は、上記端末装置において稼働しているアプリケーションを識別するアプリケーション識別情報を含み、上記提案機能は、稼働している当該アプリケーションに依存して相違し得る。かかる構成によれば、個別のアプリケーションに適合する機能がユーザに柔軟に提案されるため、より確実にユーザの不都合又は不満を解消することが可能である。

10

#### 【0083】

##### <<5. 変形例>>

本発明は上記実施形態に限定されず、様々な変形が可能である。例えば、上では何らかのアプリケーションの稼働中に起動すべき機能が提案される例を主に説明した。しかしながら、上述した実施形態は、アプリケーションの稼働中ではなく、ユーザによるOSの操作中にユーザを支援する機能を提案する目的で活用されてもよい。

#### 【0084】

また、上では学習済みモデルの例としてニューラルネットワークモデルを主に説明した。ニューラルネットワークの層の数はいくつであってもよく、多数の層を構成して所謂ディープラーニングが実行されてもよい。また、上述した実施形態は、例えば最近傍法、ナイーブベイズ法、決定木又はサポートベクターマシンといった、ニューラルネットワーク以外の機械学習アルゴリズムに基づくモデルにも適用可能である。

20

#### 【0085】

また、上ではユーザ端末の代わりに提案サーバが提案機能を決定する例を主に説明したが、ユーザ端末が自らスタンドアローン方式で提案機能を決定してもよい。その場合、ユーザ端末は、提案サーバ200の提案部230、記憶部240及び学習部250と同様の機能ブロックを有する。そして、ユーザ端末は、自ら決定した提案機能の起動の前後のユーザの表情の変化度が所定の条件を満たす場合に、提案機能の決定に用いた学習済みモデルの再学習を実行する。こうした場合にも、提案機能の決定のための学習済みモデルが実際のユーザの表情に基づく再学習を通じて強化されるため、より適切な機能をユーザに提案することが可能となる。

30

#### 【0086】

##### <<6. その他の実施形態>>

上記実施形態は、1つ以上の機能を実現するプログラムをネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける1つ以上のプロセッサがプログラムを読み出して実行する処理の形式でも実現可能である。また、1つ以上の機能を実現する回路（例えば、ASIC）によっても実現可能である。

#### 【0087】

40

発明は上記実施形態に制限されるものではなく、発明の精神及び範囲から離脱することなく、様々な変更及び変形が可能である。従って、発明の範囲を公にするために請求項を添付する。

#### 【符号の説明】

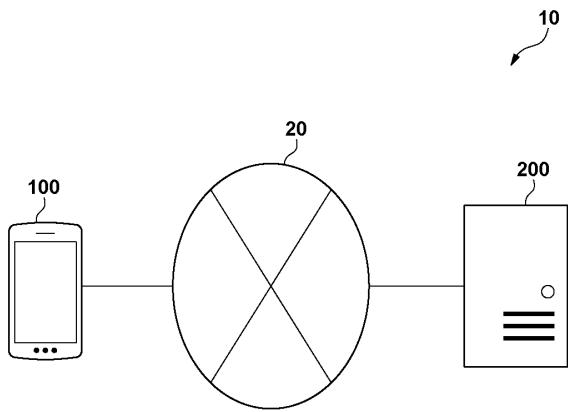
#### 【0088】

10：情報処理システム、100：ユーザ端末（端末装置）、101：コントローラ（制御手段）、111：通信I/F（通信手段）、200：提案サーバ（サーバ装置）、201：CPU（制御手段）、205：通信I/F（通信手段）、142：評価モデル、242：学習済みモデル

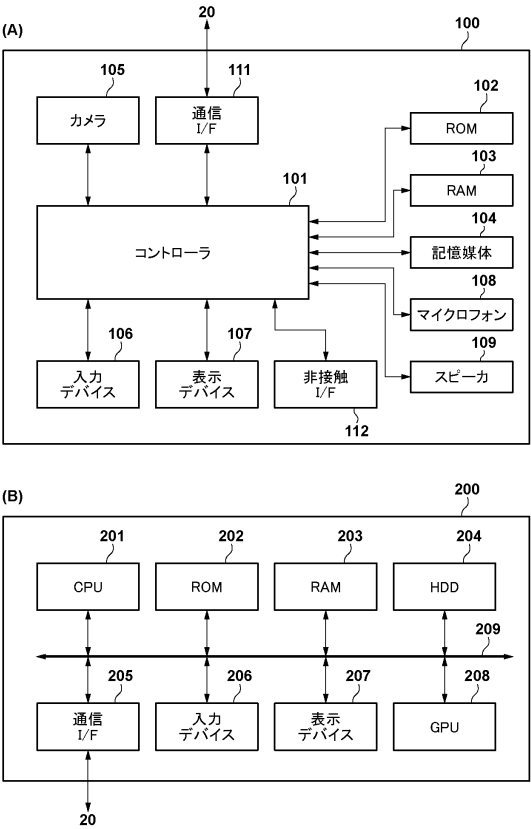
50

【図面】

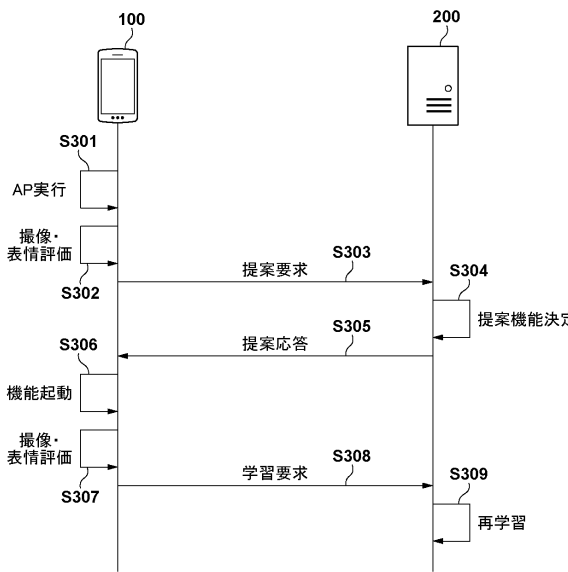
【図 1】



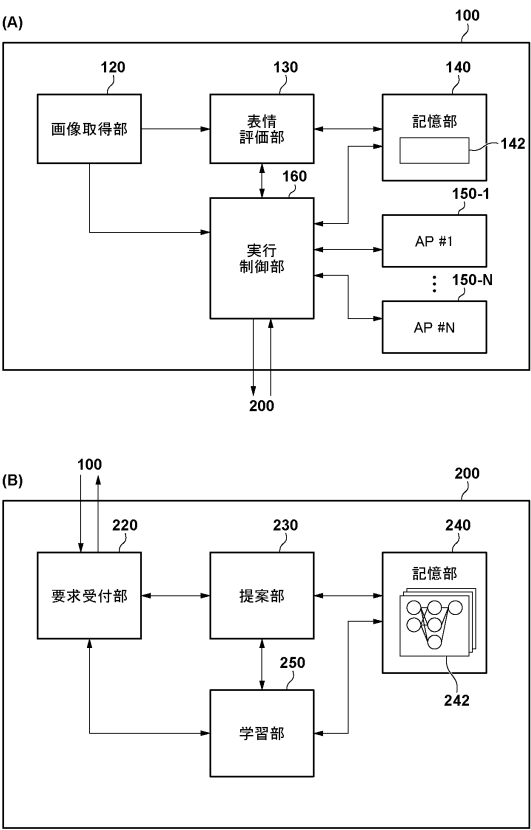
【図 2】



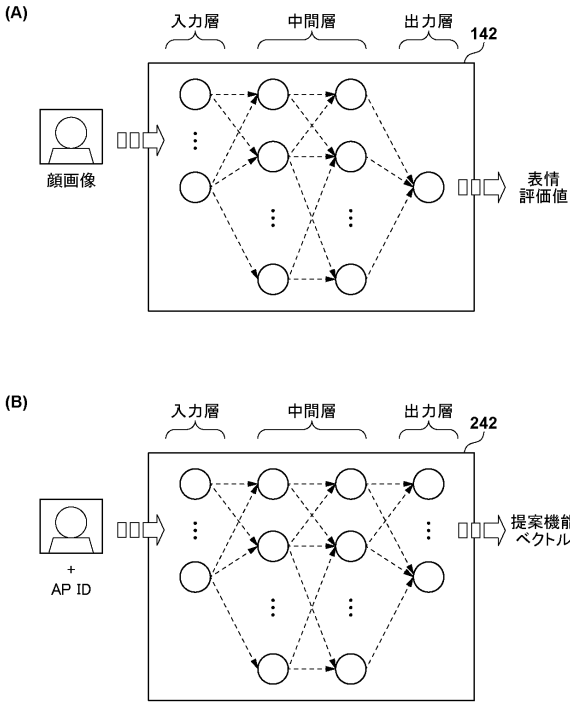
【図 3】



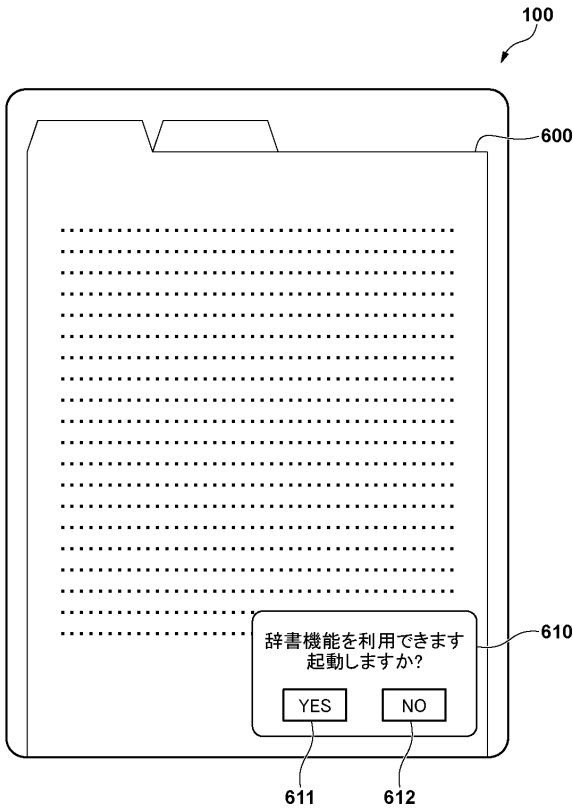
【図 4】



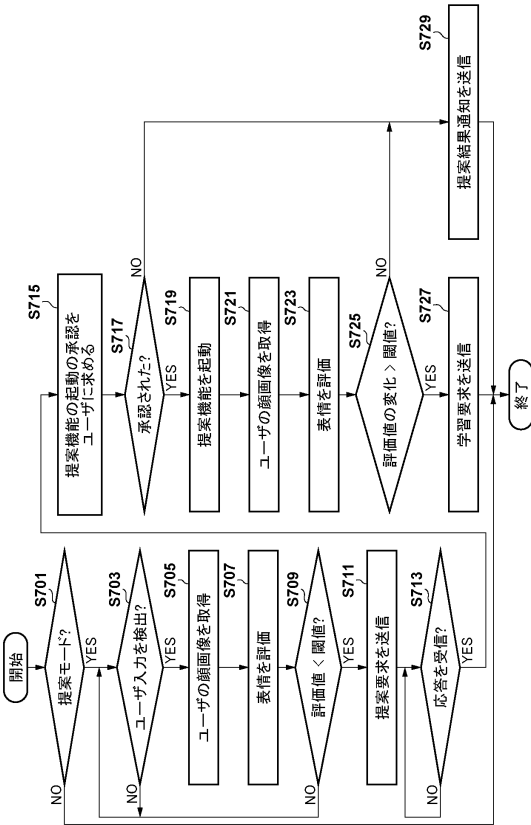
【図 5】



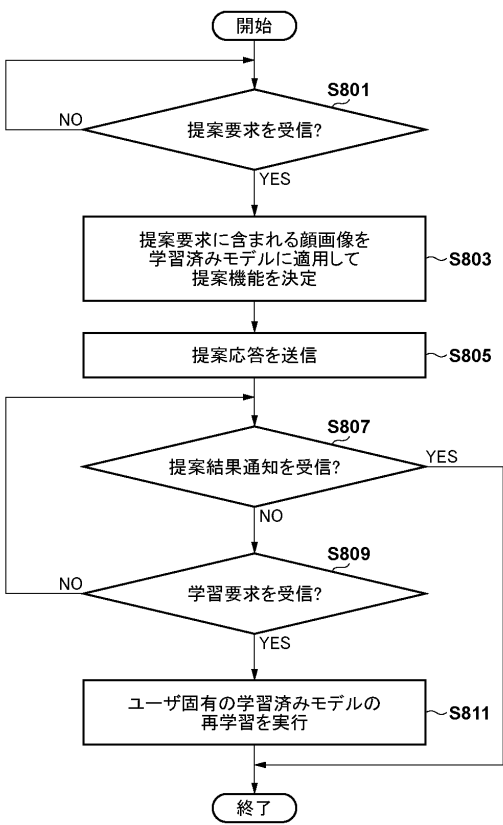
【図 6】



【図 7】



【図 8】



10

20

30

40

50

---

フロントページの続き

- (56)参考文献 特開 2 0 1 8 - 0 2 5 9 3 1 ( J P , A )  
米国特許出願公開第 2 0 1 9 / 0 2 5 0 9 3 4 ( U S , A 1 )  
特開 2 0 1 7 - 2 1 1 9 3 2 ( J P , A )  
特開 2 0 0 5 - 1 9 0 3 8 8 ( J P , A )  
特開平 0 6 - 2 6 6 4 9 7 ( J P , A )  
PARK Chang-HYun et al. , Human adaptive system model development merged with context and emotion information , 30th Annual Conference of IEEE Industrial Electronics Society, 2004. IECON 2004 , 米国 , IEEE , 2023年11月06日 , p633-636 , <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1433383> , DOI: 10.1109/IECON.2004.1433383
- (58)調査した分野 (Int.Cl. , D B 名)  
G 0 6 T 7 / 0 0  
G 0 6 F 3 / 0 1