

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7313807号
(P7313807)

(45)発行日 令和5年7月25日(2023.7.25)

(24)登録日 令和5年7月14日(2023.7.14)

(51)国際特許分類		F I	
G 0 6 F	3/16 (2006.01)	G 0 6 F	3/16 6 4 0
G 1 0 L	15/22 (2006.01)	G 1 0 L	15/22 4 5 3
G 1 0 L	17/00 (2013.01)	G 1 0 L	17/00 2 0 0 C
G 1 0 L	15/30 (2013.01)	G 1 0 L	15/30
H 0 4 M	1/00 (2006.01)	H 0 4 M	1/00 R
		請求項の数 15 (全16頁) 最終頁に続く	
(21)出願番号	特願2018-157274(P2018-157274)	(73)特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	平成30年8月24日(2018.8.24)	(74)代理人	100126240 弁理士 阿部 琢磨
(65)公開番号	特開2020-30739(P2020-30739A)	(74)代理人	100124442 弁理士 黒岩 創吾
(43)公開日	令和2年2月27日(2020.2.27)	(72)発明者	高橋 誠治 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
審査請求日	令和3年8月20日(2021.8.20)	審査官	塩屋 雅弘
		最終頁に続く	

(54)【発明の名称】 通信装置、その制御方法、およびそのプログラム

(57)【特許請求の範囲】

【請求項1】

通信装置であって、
ユーザから操作を受け付ける操作手段と、
音声を集音する集音手段と、
前記集音手段によって集音された音声を発したユーザを判断する判断手段と、
前記集音手段によって集音された音声による指示を認識する認識手段と、
外部装置から送信されるデータを受信する第一の通信手段と、
前記集音手段によって集音された音声、前記外部装置のユーザによる、前記通信装置
へのデータの送信を指示するものであると、前記判断手段および前記認識手段によって判
断された場合、前記外部装置から送信されるデータを、前記操作手段による操作を受け付
けることなく受信するよう前記第一の通信手段を制御する制御手段と、
記録媒体にユーザ情報を記録する記録手段とを有し、
前記判断手段は、前記記録手段に記録された前記ユーザ情報に基づいて、前記集音手段
によって集音された音声を発したユーザを判断する
ことを特徴とする通信装置。

【請求項2】

前記記録手段は、前記記録手段に記録された前記ユーザ情報と関連付けて前記外部装置
の情報を記録することを特徴とする請求項1に記載の通信装置。

【請求項3】

前記集音手段によって集音された音声の前記記録媒体に記録されていないユーザの声であると前記判断手段によって判断された場合、前記外部装置から前記通信装置へデータを送信することを示す通知を受信したことを、前記通信装置のユーザに報知する報知手段をさらに有する

ことを特徴とする請求項 1 または 2 に記載の通信装置。

【請求項 4】

前記報知手段は、前記集音手段によって集音された音声の前記記録媒体に記録されていないユーザの声であると前記判断手段によって判断された場合、前記外部装置からデータを受信するか否かを確認するダイアログボックスを表示部に表示する

ことを特徴とする請求項 3 に記載の通信装置。

10

【請求項 5】

前記認識手段は、前記集音手段によって集音された音声の前記記録媒体に記録されていないユーザの声であると前記判断手段によって判断され場合、前記集音手段によって集音された音声の認識を行わない

ことを特徴とする請求項 1 から 4 のいずれか 1 項に記載の通信装置。

【請求項 6】

前記ユーザ情報は声紋であることを特徴とする請求項 1 から 5 のいずれか 1 項に記載の通信装置。

【請求項 7】

前記記録手段は、前記第一の通信手段によって前記外部装置と通信したことに応じて、前記集音手段によって集音した音声と前記外部装置とを関連付けて、前記記録媒体に前記ユーザ情報として記録することを特徴とする請求項 1 から 6 のいずれか 1 項に記載の通信装置。

20

【請求項 8】

前記集音手段によって集音された音声の前記判断手段による判断処理は、前記集音手段によって集音された音声の前記認識手段による認識処理よりも先に実行される

ことを特徴とする請求項 1 から 7 のいずれか 1 項に記載の通信装置。

【請求項 9】

前記判断手段は、前記集音手段によって集音した音声に所定の単語が含まれると判断したことに応じて、前記集音手段によって集音した音声を送ったユーザを判断する

30

ことを特徴とする請求項 1 から 8 のいずれか 1 項に記載の通信装置。

【請求項 10】

前記所定の単語は、前記外部装置のユーザが所定の時間内に繰り返し発した単語であることを特徴とする請求項 9 に記載の通信装置。

【請求項 11】

前記制御手段は、前記集音手段によって集音した音声を前記第一の通信手段を介して外部サーバに送信し、

前記認識手段は、前記外部サーバから前記集音手段によって集音した音声の内容に関する情報に基づいて前記集音手段によって集音した音声の内容を認識する

ことを特徴とする請求項 1 から 10 のいずれか 1 項に記載の通信装置。

40

【請求項 12】

前記制御手段は、前記外部装置から送信されるデータを受信した場合、前記外部装置との通信を切断するよう前記第一の通信手段を制御する

ことを特徴とする請求項 1 から 11 のいずれか 1 項に記載の通信装置。

【請求項 13】

前記第一の通信手段より消費電力が低い第二の通信手段をさらに有し、

前記制御手段は、前記第一の通信手段を介して前記通信装置と接続するための接続情報を、前記第二の通信手段を介して前記外部装置へ送信し、

前記制御手段は、前記第二の通信手段を介して前記外部装置へ送信した前記接続情報を用いて、前記第一の通信手段を介して前記外部装置と接続する

50

ことを特徴とする請求項 1 から 12 のいずれか 1 項に記載の通信装置。

【請求項 14】

通信装置の制御方法であって、

ユーザから操作を受け付ける操作ステップと、

周囲の音声を集音する集音ステップと、

前記集音ステップにおいて集音された音声を発したユーザを判断する判断ステップと、

前記集音ステップにおいて集音された音声による指示を認識する認識ステップと、

前記集音ステップによって集音された音声、外部装置のユーザによるものであると前記判断ステップで判断され、かつ前記通信装置へのデータの送信を指示するものであると前記認識ステップで認識された場合、前記外部装置から送信されるデータを、前記操作ステップを行うことなく受信するステップと、

記録媒体にユーザ情報を記録するステップとを有し、

前記判断ステップでは、前記記録媒体に記録された前記ユーザ情報に基づいて、前記集音ステップで集音された音声を発したユーザを判断する

ことを特徴とする制御方法。

【請求項 15】

コンピュータを、請求項 1 乃至 13 のいずれか 1 項に記載の通信装置の各手段として機能させるための、コンピュータが読み取り可能なプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声を認識する機能を有する通信装置に関する。

【背景技術】

【0002】

通信装置同士で画像データを送受信する場合、ユーザは通信装置の画面操作を行う方法が知られている。例えば、特許文献 1 にはユーザがカメラの画面を操作して携帯電話へ画像データを送信する発明が開示されている。また、画面操作の代わりに音声入力によって通信装置を操作する方法が知られている。

【先行技術文献】

【特許文献】

【0003】

【文献】特開 2014-131225 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

音声入力を導入した場合、ユーザは画面操作の代わりに音声入力によって通信装置を操作することで、手軽に画像データの送受信することが考えられる。画像データを送信しようとするユーザは、送信する画像データを指定するため通信装置を取り出していることが想定されるため、画像データを音声入力によって送信するという操作を違和感なく行うことができる。一方、画像データを受信しようとするユーザは、画像データを受信する場合に必ずしも通信装置が手元にあるとは限らないため、画像データを音声入力によって受信するという操作に違和感を覚える場合が想定される。そこで、音声入力により送信が指示されたデータを受信する際にユーザに違和感を与えないようにすることを目的とする。

【課題を解決するための手段】

【0005】

上記目的を達成するために、本発明の通信装置は、ユーザから操作を受け付ける操作手段と、音声を集音する集音手段と、前記集音手段によって集音された音声を発したユーザを判断する判断手段と、前記集音手段によって集音された音声による指示を認識する認識手段と、外部装置から送信されるデータを受信する第一の通信手段と、前記集音手段によって集音された音声、外部装置のユーザによる、前記通信装置へのデータの送信を

10

20

30

40

50

指示するものであると前記判断手段および前記認識手段によって判断された場合、前記外部装置から送信されるデータを、前記操作手段による操作を受け付けることなく受信するよう前記第一の通信手段を制御する制御手段と、記録媒体にユーザ情報を記録する記録手段とを有し、前記判断手段は、前記記録手段に記録された前記ユーザ情報に基づいて、前記集音手段によって集音された音声を発したユーザを判断することを特徴とする。

【発明の効果】

【0006】

本発明によれば音声入力により送信が指示されたデータを受信する際にユーザに違和感を与えないようにすることができる。

【図面の簡単な説明】

【0007】

【図1】実施形態1におけるスマートフォン100のブロック図の一例。

【図2】実施形態1におけるスマートフォン100と信頼できないユーザのスマートフォンとの通信手順の一例を説明するためのシーケンス図。

【図3】実施形態1におけるスマートフォン100と信頼できるユーザのスマートフォンとの通信手順の一例を説明するためのシーケンス図。

【図4】実施形態1におけるスマートフォン100の動作の一例を示すフローチャート。

【発明を実施するための形態】

【0008】

以下に、本発明を実施するための形態について、添付の図面を用いて詳細に説明する。

【0009】

なお、以下に説明する実施の形態は、本発明の実現手段としての一例であり、本発明が適用される装置の構成や各種条件によって適宜修正又は変更されてもよい。

【0010】

〔実施形態1〕

＜各装置の構成＞

図1は、本実施形態における通信装置の一例であるスマートフォン100を示すブロック図である。これ以降の実施形態の説明ではスマートフォン100を例に挙げて説明するが、本実施形態の通信装置はスマートフォンに限定されない。本発明の通信装置は、例えばスマートフォン端末、タブレット端末、ゲーム機器、デジタルカメラ、スマートデバイス、およびスマートスピーカ等である。本実施形態の通信装置は通信機能および音声アシスタント機能を有する。音声アシスタント機能とは、集音した音声に所定の単語（いわゆるウェイクワード）が含まれる場合、そのウェイクワードに続く音声を認識し、認識した音声に対応する処理をスマートフォン100が行う機能である。例えばユーザがウェイクワードの後に「〇〇駅から△△駅までの運賃は」と発言した場合、スマートフォン100はまずユーザの音声の内容を認識するための処理を行う。この音声を認識する方法には、スマートフォン100の不揮発性メモリ103に記録している音声を認識するためのプログラムを用いる方法やスマートフォン100が外部サーバを利用する方法等がある。外部サーバを利用する場合、スマートフォン100は集音した音声を外部サーバに送信し、外部サーバにおいて認識された音声に関する情報を受信することで音声を認識する。スマートフォン100は認識した音声（発言）に対応するプログラムや外部サーバを用いて〇〇駅から△△駅までの運賃の情報を取得し、音や表示部等を用いてユーザに通知する。

【0011】

本実施形態において、スマートフォン100の所有者（ユーザ）が他の人のスマートフォン100へデータを送信する発言をする例（例えばユーザが「表示中の画像データを他のユーザに送信する」と発言する場合）を説明する。所有者は、例えばスマートフォン100に管理者（Administrator）として登録されているユーザやスマートフォン100にログイン中であるユーザである。

【0012】

制御部101は、入力された信号や、プログラムに従ってスマートフォン100の各部

10

20

30

40

50

を制御する。なお、制御部 101 が装置全体を制御する代わりに、複数のハードウェアが処理を分担することで、装置全体を制御してもよい。

【0013】

メモリ 102 は、各種データを一時的に保持するバッファメモリや、制御部 101 の作業領域等として使用される。

【0014】

不揮発性メモリ 103 は、電氣的に消去・記録可能な不揮発性のメモリであり、制御部 101 で実行されるプログラム等が格納される。このプログラムは、前述した音声アシスタント機能に関するプログラムを含む。

【0015】

操作部 104 は、スマートフォン 100 に対する指示をユーザから受け付けるために用いられる。操作部 104 は例えば、ユーザがスマートフォン 100 の電源の ON/OFF を指示するための電源ボタンや、通信機能の ON/OFF を指示するための操作ボタンを含む。また、後述する表示部 105 上に取り付けられるタッチパネルも操作部 104 に含まれる。

【0016】

表示部 105 は、対話的な操作のための GUI (Graphical User Interface) 表示を行う。なお、表示部 105 は必ずしもスマートフォン 100 が内蔵する必要はない。スマートフォン 100 は表示内容を制御する表示制御機能を少なくとも有していればよい。また、表示部 105 はタッチパネルと一体的に構成されている。例えば、タッチパネルは、光の透過率が表示部 105 の表示を妨げないように構成され、表示部 105 の表示面の上層に取り付けられる。そして、タッチパネルにおける入力座標と、表示部 105 の表示面上の表示座標とを対応付ける。これにより、あたかもユーザが表示部 105 上に表示された画面を直接的に操作可能であるかのような GUI を提供できる。

【0017】

集音部 106 は、外部の音を集音して電気信号に変換するデバイスである。集音部 106 は例えばマイクロフォン等である。

【0018】

記録媒体 107 は、各種データを記録することができる。記録媒体 107 は、スマートフォン 100 に着脱可能なように構成してもよいし、スマートフォン 100 に内蔵されていてもよい。すなわち、スマートフォン 100 は少なくとも記録媒体 107 にアクセスする手段を有していればよい。

【0019】

近距離無線通信部 111 は、例えば無線通信のためのアンテナと無線信号を処理するため変復調回路や通信コントローラから構成される。近距離無線通信部 111 は、変調した無線信号をアンテナから出力し、またアンテナで受信した無線信号を復調することにより IEEE 802.15 の無線通信規格（いわゆる Bluetooth (登録商標)）に準拠した近距離無線通信を実現する。本実施形態において Bluetooth 通信は、Bluetooth のバージョン 5.0 以降に規定されている Bluetooth Low Energy (以下、BLE と略す) という仕様を採用する。BLE 通信は、後述する無線通信部 112 よりも低消費電力である。制御部 101 は、近距離無線通信部 111 を制御し、他の機器との BLE 通信を実現する。

【0020】

無線通信部 112 は、無線通信を実現するための通信ユニットである。無線通信部 112 は、無線通信のためのアンテナと無線信号を処理するための通信コントローラから構成される。本実施形態において無線通信部 112 は、IEEE 802.11 の無線通信規格に準拠した無線 LAN 通信、いわゆる Wi-Fi (登録商標) で通信するためのインターフェースを含む。制御部 101 は、無線通信部 112 を制御し、他の機器との無線 LAN 通信を実現する。本実施形態において無線 LAN 通信は BLE 通信よりも通信速度が速い。

【0021】

10

20

30

40

50

なお、本実施形態においてスマートフォン１００はウェイクワードや後述する信頼できるユーザの音声を記録するために、電源ＯＮの状態において音声を集音している。電源ＯＮの状態は、例えば、制御部１０１、メモリ１０２、不揮発性メモリ１０３、集音部１０６、記録媒体１０７、近距離無線通信部１１１、および無線通信部１１２が動作している状態である。この状態は、表示部１０５等がＯＦＦであるような省電力な状態も含む。

【００２２】

また、本実施形態においてスマートフォン１００には信頼できるユーザおよび信頼できるユーザの音声を登録できる。信頼できるユーザは、例えばユーザがスマートフォン１００を操作して電話帳やアカウント等に登録されているユーザの中から信頼できるとしたユーザのことである。なお、所有者は信頼できるユーザに含まれる。信頼できるユーザの音声は、例えば信頼できるユーザに関連付けて登録された音声、所定の時間内に認識した音声、および所定の音量以上の音声等である。スマートフォン１００は信頼できるユーザの音声の特徴である声紋をユーザ情報としてメモリ１０２、不揮発性メモリ１０３、および記録媒体１０７の少なくとも一つに記録している。スマートフォン１００に音声を登録する方法には、例えばユーザが友人の音声を集音して、その音声をその友人のアカウントに登録する方法がある。所定の時間内に認識した音声は、例えばウェイクワードを集音する５分前までに認識した音声のことである。なおスマートフォン１００は現在から５分以内における音声を記録しているが、５分より前の音声については削除する。所定の時間内に認識した音声および所定の音量以上の音声を信頼できるユーザとする理由は、例えばユーザ同士が歓談中にデータの送受信を行うユースケースを想定している。その場合においては所定の時間内または所定の音量以上の音声が発せられていると考えられるからである。

【００２３】

<システムの概要>

図２および図３を用いて、本実施形態における通信装置の一例を説明する。

【００２４】

図２はスマートフォンＡ１００とスマートフォンＢ１００との間の無線通信の一例を示したシーケンス図である。図３は、スマートフォンＡ１００とスマートフォンＣ１００との間の無線通信の一例を示したシーケンス図である。ここで、ユーザＡはスマートフォンＡ１００に所有者として登録されている。ユーザＢはスマートフォンＢ１００に所有者として登録されている。ユーザＣはスマートフォンＣ１００に所有者として登録されている。ユーザＡ、ユーザＢ、およびユーザＣはそれぞれ異なる人物である。ユーザＡはスマートフォンＡ１００以外のスマートフォンには所有者として登録されていない。ユーザＢはスマートフォンＢ１００以外のスマートフォンには所有者として登録されていない。ユーザＣはスマートフォンＣ１００以外のスマートフォンには所有者として登録されていない。また本実施形態において、スマートフォンＢ１００ではユーザＡは信頼できるユーザとして記録されておらず、スマートフォンＣ１００ではユーザＡは信頼できるユーザとして記録されているものとする。例えば、ユーザＡはユーザＣとは友達であるがユーザＢとは初対面の状況において、ユーザＡがスマートフォンＡ１００を用いてユーザＢおよびユーザＣを撮影し、ユーザＡは音声での指示によってその画像データをユーザＢおよびユーザＣに送ることを想定する。ただし、例えばユーザＡとユーザＢが歓談した場合、スマートフォンＢ１００はユーザＡを信頼できるユーザとして記録するが、説明を簡易にするためにそのような場合を本実施形態では考慮しない。

【００２５】

図２および図３において、スマートフォンＡ１００、スマートフォンＢ１００、およびスマートフォンＣ１００は、いずれもスマートフォン１００と同様の構成を有する。また図２および図３において、ＢＬＥ通信は破線の矢印、無線ＬＡＮ通信は実線の矢印で示す。図２および図３において、ユーザＡが、スマートフォンＢ１００またはスマートフォンＣ１００に対して画像データを送信する指示を音声によって行った場合を例に説明する。具体的には、ユーザＡがウェイクワードに続いて「表示中の画像データをユーザＢ（またはユーザＣ）に送る」等とスマートフォンＡ１００に言った場合である。図２において、

ユーザAがスマートフォンA100からスマートフォンB100へ画像データを送信する例を説明する。スマートフォンA100およびスマートフォンB100はウェイクワードを集音したことをトリガに本シーケンスにおける処理を開始する。

【0026】

ステップS201において、ユーザAは音声によってスマートフォンA100からユーザBへ画像データの送信を指示する。例えば、ユーザAがウェイクワードに続いて「表示中の画像データをユーザBに送る」と指示する。ステップS202において、ユーザAはステップS201と同様に指示する。ここで、ステップS201、S202においてスマートフォンA100およびスマートフォンB100はユーザAから個別に指示されたのではなく、同じ指示（音声）を集音している。言いかえると、ユーザAはスマートフォンA100、およびスマートフォンB100のそれぞれに指示しているのではなく、1回だけ指示している。この時点では、スマートフォンA100およびスマートフォンB100は音声の内容について認識するための処理を行わない。

10

【0027】

ステップS203において、スマートフォンA100は、ステップS201において集音した音声信頼できるユーザか否かを判断する。例えばスマートフォンA100はメモリ102、不揮発性メモリ103、または記録媒体107に記録されている声紋と集音した音声の声紋とを比較して、2つの声紋が一致するか否かを判断する。本実施形態において、スマートフォンA100はステップS201において集音した音声信頼できるユーザAの音声と一致すると判断し、ステップS205の処理を行う。スマートフォンA100は、ステップS204において、スマートフォンB100はステップS203と同様に、集音した音声信頼できるユーザを判断する。本実施形態において、スマートフォンB100はステップS202において集音した音声信頼できるユーザAの声紋と一致しないと判断し、音声の内容を認識するための処理を行わない。

20

【0028】

ステップS205において、スマートフォンA100はステップS201において集音した音声の内容を認識する。例えばスマートフォンA100は集音した音声を外部サーバに送信し、外部サーバにおいて認識された音声に関する情報を受信することで音声信頼できるユーザAであることを認識する。本実施形態においてこの情報は、「スマートフォンA100が現在表示している画像をユーザBに送信する」ことを意味する情報である。

30

【0029】

ステップS206において、スマートフォンA100はステップS205の認識に基づき、画像データを送信することを示す通知をBLE通信によりブロードキャストする。この場合、スマートフォンA100は、ブロードキャストした通知にユーザBを示す情報（例えば、ユーザ名やスマートフォンB100のデバイスID等）を含める。スマートフォンB100はこの通知を受信し、所有者であるユーザBへの画像データの送信であると判断し、スマートフォンA100から画像データを受信するための処理を行う。

【0030】

ステップS207において、スマートフォンB100は、ステップS204においてユーザAを信頼できないユーザと判断したため、スマートフォンA100から通知が来たことをスマートフォンB100の所有者であるユーザBに報知する。例えば、スマートフォンB100は表示部105にスマートフォンA100から画像データを受信するか否かを確認するダイアログボックスを表示する。

40

【0031】

ステップS208において、ユーザBはスマートフォンB100にスマートフォンA100から画像データを受信するように操作を行う。例えば、ユーザBがステップS207において表示されたダイアログボックスにあるYESボタンを押す等の操作がある。

【0032】

ステップS209において、スマートフォンB100はスマートフォンA100にBLE通信によって画像データを受信することを示す通知を送信する。スマートフォンB100

50

0 は、少なくともスマートフォン A 1 0 0 との無線 LAN 接続に必要な接続情報（例えば、SSID や IP アドレス等）をこの通知に含める。

【0033】

ステップ S 2 1 0 において、スマートフォン A 1 0 0 とスマートフォン B 1 0 0 は、無線 LAN 接続する。スマートフォン A 1 0 0 はこの無線 LAN 接続において、ステップ S 2 0 9 においてスマートフォン B 1 0 0 から受信した情報を用いる。

【0034】

ステップ S 2 1 1 において、スマートフォン A 1 0 0 は、スマートフォン B 1 0 0 に対して、無線 LAN 通信によって画像データを送信する。この画像データは、ステップ S 2 0 1 においてユーザがスマートフォン A 1 0 0 に対してスマートフォン B 1 0 0 へ送信するように指示した画像データである。

10

【0035】

ステップ S 2 1 2 において、スマートフォン A 1 0 0 は、スマートフォン B 1 0 0 との無線 LAN 接続を切断する。

【0036】

以上、図 2 において、スマートフォン A 1 0 0 とスマートフォン B 1 0 0 との間の無線通信の一例を説明した。

【0037】

なお、ステップ S 2 0 8 において、ユーザ B がスマートフォン A 1 0 0 から画像データを受信しないように操作した場合、ステップ S 2 0 9 においてスマートフォン B 1 0 0 はスマートフォン A 1 0 0 へ画像データを受信しないことを示す通知を送信する。例えば、ステップ S 2 0 8 ユーザ B がステップ S 2 0 7 において表示されたダイアログボックスにある NO ボタンを押す等の操作によって、スマートフォン B 1 0 0 はスマートフォン A 1 0 0 へ画像データを受信しない通知を送信する。

20

【0038】

次に図 3 においてスマートフォン A 1 0 0 とスマートフォン C 1 0 0 との間の無線通信の一例を説明する。図 3 において、ユーザ A がスマートフォン A 1 0 0 からスマートフォン C 1 0 0 へ画像データを送信する例を説明する。スマートフォン A 1 0 0 およびスマートフォン C 1 0 0 はウェイクワードを集音したことをトリガに本シーケンスにおける処理を開始する。ここでスマートフォン A 1 0 0 およびスマートフォン C 1 0 0 はユーザ A を信頼できるユーザとして認識する。

30

【0039】

ステップ S 3 0 1 において、ユーザ A は音声によってスマートフォン A 1 0 0 からユーザ C へ画像データの送信を指示する。例えば、ユーザ A がウェイクワードに続いて「表示中の画像データをユーザ C に送る」と指示する。ステップ S 3 0 2 において、ユーザ A はステップ S 3 0 1 と同様に指示する。ここで、ステップ S 3 0 1、S 3 0 2 において、スマートフォン A 1 0 0 およびスマートフォン C 1 0 0 はユーザ A から個別に指示されたのではなく、同じ指示（音声）を集音している。言いかえると、ユーザ A はスマートフォン A 1 0 0、およびスマートフォン C 1 0 0 のそれぞれに指示しているのではなく、1 回だけ指示している。この時点では、スマートフォン A 1 0 0 およびスマートフォン C 1 0 0 は音声の内容について認識するための処理を行わない。

40

【0040】

ステップ S 3 0 3 において、スマートフォン A 1 0 0 は、ステップ S 3 0 1 において集音した音声信頼できるユーザか否かを判断する。例えばスマートフォン A 1 0 0 はメモリ 1 0 2、不揮発性メモリ 1 0 3、または記録媒体 1 0 7 に記録されている声紋と集音した音声の声紋とを比較して、2 つの声紋が一致するか否かを判断する。本実施形態において、スマートフォン A 1 0 0 はステップ S 3 0 1 において集音した音声信頼できるユーザ A の所有者（ユーザ A）の音声と一致すると判断し、ステップ S 3 0 5 の処理を行う。ステップ S 3 0 4 において、スマートフォン C 1 0 0 はステップ S 3 0 3 と同様に、集音した音声を発したユーザを判断する。本実施形態において、スマートフォン C 1 0 0 は

50

ステップ S 3 0 2 において集音した音声の声紋はスマートフォン C 1 0 0 の所有者ではないが信頼できるユーザ（ユーザ A）の声紋と一致すると判断し、ステップ S 3 0 6 の処理を行う。

【 0 0 4 1 】

ステップ S 3 0 5 において、スマートフォン A 1 0 0 はステップ S 3 0 1 において集音した音声の内容を認識する。例えばスマートフォン A 1 0 0 は集音した音声を外部サーバに送信し、外部サーバにおいて認識された音声に関する情報を受信することで音声を認識する。本実施形態においてこの情報は、「スマートフォン A 1 0 0 が現在表示している画像をユーザ C に送信する」ことを意味する情報である。ステップ S 3 0 6 において、スマートフォン C 1 0 0 はステップ S 3 0 5 と同様に、集音した音声の内容を認識する処理を行う。

10

【 0 0 4 2 】

ステップ S 3 0 7 において、スマートフォン A 1 0 0 はステップ S 3 0 5 の認識に基づき、画像データを送信することを示す通知を BLE 通信によってブロードキャストする。この場合、スマートフォン A 1 0 0 は、ブロードキャストした通知にユーザ C を示す情報（例えば、ユーザ名やスマートフォン C 1 0 0 のデバイス ID 等）を含める。スマートフォン C 1 0 0 はこの通知を受信し、所有者であるユーザ C への画像データの送信であると判断し、スマートフォン A 1 0 0 から画像データを受信するための処理を行う。

【 0 0 4 3 】

ステップ S 3 0 8 において、スマートフォン C 1 0 0 は、ステップ S 3 0 4 においてユーザ A を信頼できるユーザと判断したため、スマートフォン A 1 0 0 に BLE 通信によって画像データを受信することを示す通知を自動で送信する。スマートフォン C 1 0 0 は、少なくともスマートフォン A 1 0 0 との無線 LAN 接続に必要な接続情報（例えば、SSID や IP アドレス等）をこの通知に含める。

20

【 0 0 4 4 】

ステップ S 3 0 9 において、スマートフォン A 1 0 0 とスマートフォン C 1 0 0 は、無線 LAN 接続する。スマートフォン A 1 0 0 はこの無線 LAN 接続において、ステップ S 2 0 9 においてスマートフォン C 1 0 0 から受信した情報を用いる。

【 0 0 4 5 】

ステップ S 3 1 0 において、スマートフォン A 1 0 0 は、スマートフォン C 1 0 0 に対して、無線 LAN 通信によって画像データを送信する。この画像データは、ステップ S 3 0 1 においてユーザがスマートフォン A 1 0 0 に対してスマートフォン C 1 0 0 へ送信するように指示した画像データである。

30

【 0 0 4 6 】

ステップ S 3 1 1 において、スマートフォン A 1 0 0 は、スマートフォン C 1 0 0 との無線 LAN 接続を切断する。

【 0 0 4 7 】

以上、図 3 において、スマートフォン A 1 0 0 とスマートフォン C 1 0 0 との間の無線通信の一例を説明した。

【 0 0 4 8 】

40

< スマートフォン 1 0 0 の動作 >

続いて、上記の動作を実現するためのスマートフォン 1 0 0 の処理手順について、図 4 を参照しながら説明する。

【 0 0 4 9 】

図 4 は、本実施形態におけるスマートフォン 1 0 0 の処理手順を示すフローチャートの一例である。なお、本フローチャートに示す処理は、スマートフォン 1 0 0 の制御部 1 0 1 がユーザによって入力された信号や不揮発性メモリ 1 0 3 等に記録されたプログラムに従い、スマートフォン 1 0 0 の各部を制御することにより実現される。

【 0 0 5 0 】

本フローチャートは、音声アシスタント機能が有効であるスマートフォン 1 0 0 が電源

50

ONの状態において、制御部101によって実行される。

【0051】

ステップS400において、制御部101は集音部106を用いて集音した音声にウェイクワードが含まれているか否かを判断する。ウェイクワードが含まれている場合、制御部101はステップS401の処理を行う。ウェイクワードが含まれていない場合、制御部101はステップS418の処理を行う。

【0052】

ステップS401において、制御部101は集音部106を用いて集音した音声信頼できるユーザによるものか否かを判断する。制御部101はユーザの判断を、不揮発性メモリ103に記録されているプログラムを用いて行う。例えば、制御部101はメモリ102または不揮発性メモリ103に記録されている信頼できるユーザの声紋と集音した音声の声紋とが一致するか否かを判断する。制御部101は音声信頼できるユーザであると判断した場合は、ステップS402の処理を行う。制御部101は音声信頼できるユーザではないと判断した場合は、ステップS418の処理を行う。本ステップは図2のステップS203、ステップS204および図3のステップS303およびステップS304に対応する。

【0053】

ステップS402において、制御部101は音声アシスタント機能を用いて音声を認識し、画像データを送信するか受信するかを判断する。例えば制御部101は、集音した音声を外部サーバに送信し、外部サーバにおいて認識された音声に関する情報を受信することで音声を認識する。例えば、ユーザAが「表示中の画像データをユーザCに送る」と発言した場合、スマートフォンC100は自身の所有者であるユーザCに送信されると判断できるため、スマートフォンC100の制御部101は画像データを受信すると判断する。また、この場合、この発言の中にスマートフォンA100または自身の所有者であるユーザAを示す情報はないため、スマートフォンA100の制御部101は画像データを送信すると判断する。なお、スマートフォンB100はステップS401においてユーザAを信頼できないユーザと判断するため、スマートフォンB100の制御部101は本ステップの処理を行わない。データを送信すると判断した場合、ステップS420の処理を行う。制御部101は画像データを受信すると判断した場合、ステップS412の処理を行う。本ステップは図2のステップS205および図3のステップS305、ステップS306に対応する。

【0054】

まず、スマートフォン100がデータを送信する場合について説明する。例えば、制御部101は他のユーザへ画像データを送信する指示を集音した場合にこの処理を行う。

【0055】

ステップS403において、制御部101は画像データを送信することを示す通知を近距離無線通信部111を用いてブロードキャストする。制御部101はブロードキャストした通知に、他のユーザを示す情報（例えば、ユーザ名やデバイスID等）を含める。本ステップは図2のステップS206および図3のステップS307に対応する。

【0056】

ステップS404において、制御部101はステップS403においてブロードキャストした通知に対する応答を近距離無線通信部111によって受信したか否かを判断する。制御部101が該応答を受信した場合はステップS405の処理を行う。制御部101が該応答を受信していない場合はステップS410の処理を行い、所定の時間が経過しているか否かを判断する。本ステップは図2のステップS209および図3のステップS308に対応する。

【0057】

ステップS405において、制御部101はステップS404において受信した応答の内容が他のスマートフォンが画像データを受信することを示すか否かを判断する。他のスマートフォンが画像データを受信することを示す応答だった場合、制御部101はステッ

10

20

30

40

50

プ S 4 0 6 の処理を行う。他のスマートフォンが画像データを受信しないことを示す応答だった場合、制御部 1 0 1 はステップ S 4 0 9 の処理を行う。

【 0 0 5 8 】

ステップ S 4 0 6 において、制御部 1 0 1 はステップ S 4 0 4 において受信した応答に含まれる無線 LAN 接続に必要な情報を用いて、無線通信部 1 1 2 によって他のスマートフォンと無線 LAN 接続する。本ステップは図 2 のステップ S 2 1 0 および図 3 のステップ S 3 0 9 に対応する。

【 0 0 5 9 】

ステップ S 4 0 7 において、制御部 1 0 1 は他のスマートフォンと無線通信部 1 1 2 を用いた無線 LAN 通信によって、ユーザに送信を指示された画像データを送信する。本ステップは図 2 のステップ S 2 1 1 および図 3 のステップ S 3 1 0 に対応する。

【 0 0 6 0 】

ステップ S 4 0 8 において、制御部 1 0 1 は他のスマートフォンとの無線 LAN 接続を切断する。本ステップは図 2 のステップ S 2 1 2 および図 3 のステップ S 3 1 1 に対応する。

【 0 0 6 1 】

ステップ S 4 0 9 において、制御部 1 0 1 は他のスマートフォンから画像データの受信を拒否されたことをユーザに報知して処理を終了する。例えば制御部 1 0 1 は、表示部 1 0 5 に「画像データの受信が拒否されました」等のメッセージを表示することやスピーカ（不図示）等を用いてメッセージを読み上げる音声を出力することでユーザに報知する。

【 0 0 6 2 】

ステップ S 4 1 0 において、制御部 1 0 1 はステップ S 4 0 3 においてブロードキャストしてから所定の時間以上経過したか否か判断する。制御部 1 0 1 は所定の時間が経過していないと判断した場合、ステップ S 4 0 4 の処理を行い、他のスマートフォンから通知を受信するまで待機する。制御部 1 0 1 は所定の時間が経過したと判断した場合、ステップ S 4 1 1 の処理を行う。

【 0 0 6 3 】

ステップ S 4 1 1 において、制御部 1 0 1 は他のスマートフォンから応答を受信できなかったことをユーザに報知して処理を終了する。例えば制御部 1 0 1 は、表示部 1 0 5 に「画像データの送信ができませんでした」等のメッセージを表示することやスピーカ（不図示）等を用いてメッセージを読み上げる音声を出力することでユーザに報知する。これによりユーザは画像データの送信に失敗したことに気付いてすぐに再送を試みることができる。また、制御部 1 0 1 はステップ S 4 0 9 と異なるメッセージを表示することで、エラーの種類を判別できる。

【 0 0 6 4 】

ここまで、スマートフォン 1 0 0 が画像データを送信する場合について説明した。次にスマートフォン 1 0 0 が信頼できるユーザから画像データを受信する場合について説明する。

【 0 0 6 5 】

ステップ S 4 1 2 において、制御部 1 0 1 はスマートフォン 1 0 0 に対して画像データを送信することを示す通知を近距離無線通信部 1 1 1 によって受信したか否かを判断する。制御部 1 0 1 がその通知を受信した場合はステップ S 4 1 3 の処理を行う。制御部 1 0 1 がその通知を受信していない場合はステップ S 4 1 7 の処理を行い、所定の時間が経過しているか否か判断する。本ステップは図 3 のステップ S 3 0 7 に対応する。

【 0 0 6 6 】

ステップ S 4 1 3 において、制御部 1 0 1 はステップ S 4 1 2 において受信した通知の送信元となるスマートフォンに対して、自動で近距離無線通信部 1 1 1 によって、画像データを受信することを示す通知を送信する。制御部 1 0 1 はこの通知に無線 LAN 接続に必要な接続情報（例えば、SSID や IP アドレス等）を含める。このように信頼できるユーザに対しては自身の所有者でなくとも自動で画像データを受信する処理を行うことで

10

20

30

40

50

、スマートフォン１００のユーザは画像データを受信する操作を行わなくて済む。本ステップは図３のステップＳ３０８に対応する。

【００６７】

ステップＳ４１４において、制御部１０１はステップＳ４１３において送信した通知に含めた無線ＬＡＮ接続に必要な情報を用いて、無線通信部１１２によって他のスマートフォンと無線ＬＡＮ接続する。本ステップは図３のステップＳ３０９に対応する。

【００６８】

ステップＳ４１５において、制御部１０１は他のスマートフォンと無線通信部１１２を用いた無線ＬＡＮ通信によって画像データを受信する。本ステップは図３のステップＳ３１０に対応する。

【００６９】

ステップＳ４１６において、制御部１０１は他のスマートフォンとの無線ＬＡＮ接続を切断する。本ステップは図３のステップＳ３１１に対応する。

【００７０】

ステップＳ４１７において、制御部１０１はステップＳ４０２において音声を認識してから所定の時間以上経過したか否か判断する。制御部１０１は所定の時間が経過していないと判断した場合、ステップＳ４１２の処理を行い、他のスマートフォンから通知を受信するまで待機する。制御部１０１は所定の時間が経過したと判断した場合、処理を終了する。ここで制御部１０１は所定の時間が経過してもユーザに報知をしない理由は、ユーザに受信する意思がないと判断できるからである。

【００７１】

ここまで、制御部１０１がステップＳ４００においてウェイクワードを認識した場合およびステップＳ４０１において集音した音声信頼できるユーザと判断した場合について説明した。次に、制御部１０１がステップＳ４００においてウェイクワードを認識しない場合またはステップＳ４０１において集音した音声信頼できないユーザと判断した場合について説明する。

【００７２】

ステップＳ４１８において、制御部１０１はスマートフォン１００に対して画像データを送信することを示す通知を近距離無線通信部１１１によって受信したか否かを判断する。制御部１０１がその通知を受信した場合はステップＳ４１９の処理を行う。制御部１０１がその通知を受信していない場合はステップＳ４００の処理に戻る。本ステップは図２のステップＳ２０６に対応する。

【００７３】

ステップＳ４１９において、制御部１０１はユーザに画像データを送信することを示す通知を受信したことをユーザに報知する。例えば制御部１０１は表示部１０５に画像を受信するか否かを確認するダイアログボックスを表示する。ユーザはこのダイアログボックスに対してタッチ操作によって受信するか否かの操作を行う。本ステップは図２のステップＳ２０７に相当する。

【００７４】

ステップＳ４２０において、制御部１０１はユーザが操作した内容が画像データを受信するための操作か受信しないための操作かを判断する。制御部１０１は画像データを受信するための操作と判断した場合、ステップＳ４１３の処理を行い、画像データを受信するための処理を行う。制御部１０１は画像を受信しないための操作と判断した場合、ステップＳ４２１において画像データを受信しないための処理を行う。

【００７５】

ステップＳ４２１において、制御部１０１は画像データを送信する通知を送信したスマートフォンに対して、画像を受信しないことを示す通知を近距離無線通信部１１１によって送信する。

【００７６】

以上、スマートフォン１００の詳細な動作について、図４を参照しながら説明した。

10

20

30

40

50

【0077】

また、ステップS417において所定の時間が経過した後に制御部101は通知を表示せずに処理を終了するが、ステップS410と同様に所定の時間が経過した後に制御部101は通知を表示してもよい。例えば制御部101は、表示部105に「画像データの受信ができませんでした」や「Bluetoothの機能がOFFです」等の通知を表示する、またはスピーカ（不図示）等を用いてユーザに音で報知する等の処理を行う。制御部101がこのような処理を行うことで、ユーザはスマートフォン100が画像データを受信できない状態にあることに気付くことができる。

【0078】

以上のように、本実施形態におけるスマートフォン100は、ユーザ操作を省略してデータを受信できる。これによって、例えばユーザがスマートフォン100をポケットに入れたまま画像データを受信できる。また、信頼できるユーザによるものでないスマートフォンからデータを受信する場合にはユーザ操作を必要とすることで、意図しないユーザから画像データを勝手に受信しないようにできる。

【0079】

＜その他の実施形態＞

上述の実施形態では、スマートフォン100に記録されている画像データを送信する場合を例に挙げて説明したが、受信するデータはこれに限定されず、動画データや音声データであってもよい。また、スマートフォン100は特定のサービスのコンテンツ（例えば、電子マネーや絵文字等のデータ）を受信してもよい。ここでいう絵文字は電子的なデータである。絵文字は例えばSNSのチャットで用いられる。

【0080】

上述の実施形態におけるスマートフォン100は、ウェイクワードをユーザごとに関連付けて記録する構成でもよい。例えば、特定のユーザが同じ単語を所定の時間内に所定回数発していると判定した場合、スマートフォン100はその単語を当該ユーザのウェイクワードとして判断し、そのユーザの声紋と一緒にユーザ情報として記録する。スマートフォン100はユーザごとにウェイクワードと声紋とを用いることで信頼できるユーザの音声他ユーザの音声と判別しやすくなる。

【0081】

上述の実施形態におけるスマートフォン100は、ユーザの声紋に加えて、顔情報に基づいて信頼できるか否かを判断してもよい。例えば、上述の実施形態におけるスマートフォン100が不図示の撮像手段を有しており、音声の集音と並行して周囲の人物の顔を撮像する。音声の集音と撮像した人物の口の動きのタイミングが一致しており、且つ、撮像した人物の顔を信頼できるユーザとしてユーザ情報に登録済みである場合、スマートフォン100は集音した音声を信頼できるユーザによるものと判定する。これにより、信頼できないユーザからデータを勝手に受信してしまう可能性を減らすことができる。

【0082】

上述の実施形態におけるスマートフォン100は、ユーザの声紋に加えて、所定の時間以内にパケットを受信したことのある他のスマートフォンを信頼できるユーザのスマートフォンとして判断してもよい。例えば、他のスマートフォンは識別情報（例えば、デバイスID等）を繰り返しブロードキャストし、スマートフォン100はその識別情報を受信する。スマートフォン100はその識別情報を送信したスマートフォンと同じタイミングで集音した音声の声紋とを関連付けてユーザ情報として記録する。信頼できるユーザとスマートフォンを関連付けることで、より確実に受信したい画像データを受信できる。

【0083】

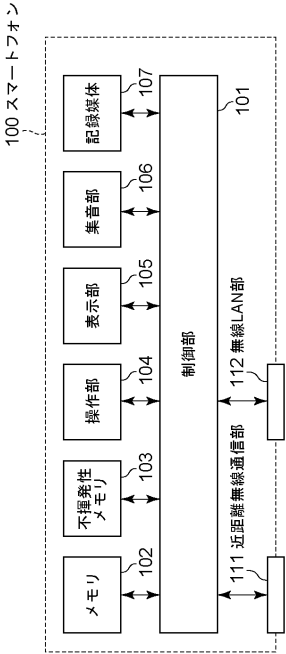
本発明は、上述の実施形態の1以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける1つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1以上の機能を実現する回路（例えば、ASIC）によっても実現可能である。

【0084】

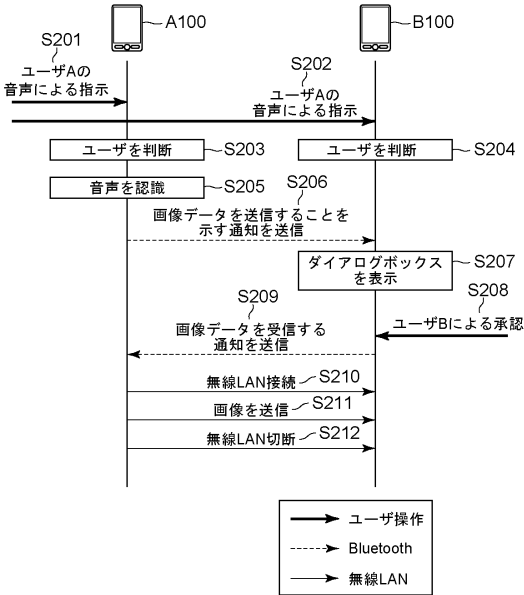
なお、本発明は上記実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、上記実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせてもよい。

【図面】

【図 1】



【図 2】



10

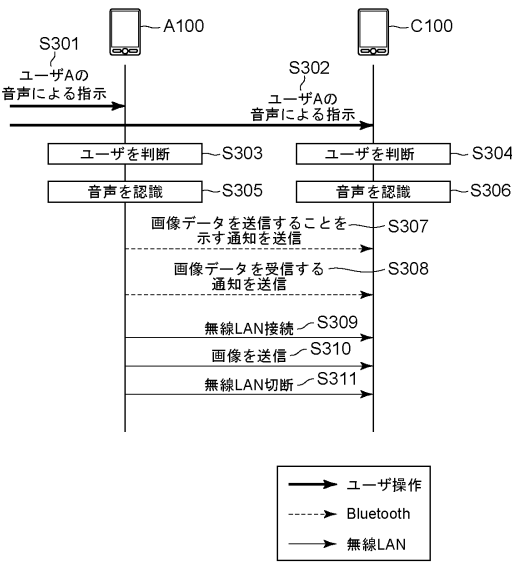
20

30

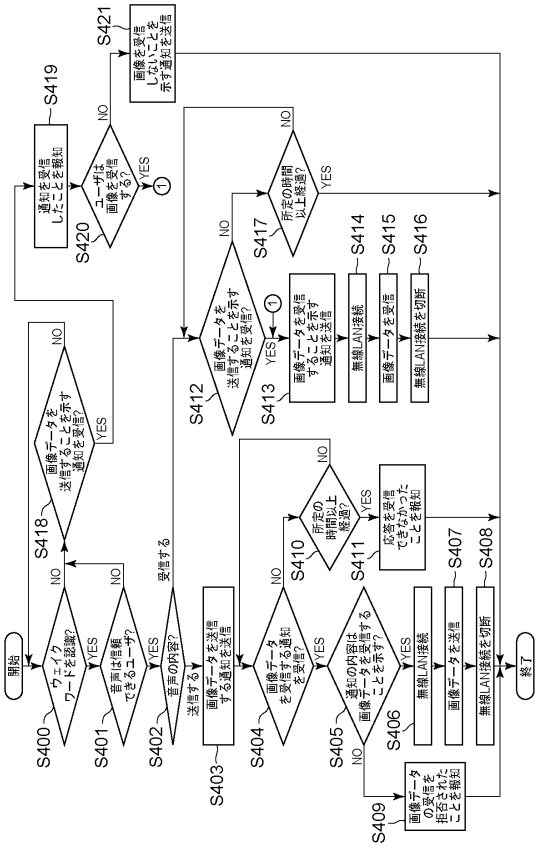
40

50

【図 3】



【図 4】



10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I
G 0 6 F 3/16 6 5 0

(56)参考文献

特開 2 0 1 7 - 1 6 7 6 8 5 (J P , A)

米国特許出願公開第 2 0 1 4 / 0 0 1 2 5 8 7 (U S , A 1)

(58)調査した分野 (Int.Cl., D B 名)

G 0 6 F 3 / 0 1

3 / 0 4 8 - 3 / 0 4 8 9 5

3 / 1 6

G 1 0 L 1 5 / 0 0 - 1 7 / 2 6

H 0 4 L 5 1 / 0 0 - 5 1 / 5 8

6 7 / 0 0 - 6 7 / 7 5

H 0 4 M 1 / 0 0

1 / 2 4 - 1 / 8 2

9 9 / 0 0