

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-243392

(P2007-243392A)

(43) 公開日 平成19年9月20日(2007.9.20)

(51) Int. Cl.	F I	テーマコード (参考)
HO4M 3/56 (2006.01)	HO4M 3/56 Z	5D015
G1OL 15/00 (2006.01)	G1OL 15/00 200A	5K201
G1OL 15/22 (2006.01)	G1OL 15/22 200H	
G1OL 13/00 (2006.01)	G1OL 13/00 100N	
G1OL 13/02 (2006.01)	G1OL 13/02 130Z	

審査請求 未請求 請求項の数 19 OL (全 18 頁) 最終頁に続く

(21) 出願番号 特願2006-60764 (P2006-60764)
 (22) 出願日 平成18年3月7日(2006.3.7)

(71) 出願人 000004237
 日本電気株式会社
 東京都港区芝五丁目7番1号
 (74) 代理人 100124811
 弁理士 馬場 資博
 (72) 発明者 深作 正樹
 東京都港区芝五丁目7番1号 日本電気株式会社内
 (72) 発明者 三上 典昭
 東京都港区芝五丁目7番1号 日本電気株式会社内
 Fターム(参考) 5D015 KK02 KK04
 5K201 AA05 BB09 BC23 BC27 BD06
 CA02 DC05 EA07 ED05 EF03
 EF06

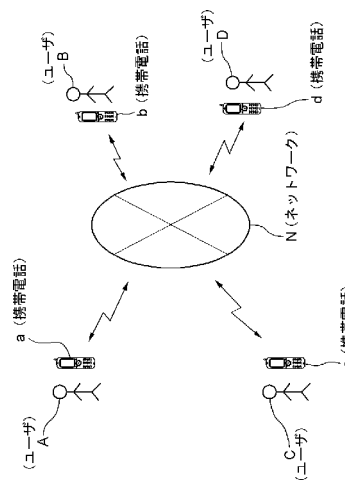
(54) 【発明の名称】 通話端末、多者通話システム、多者通話方法、プログラム

(57) 【要約】

【課題】グループセッション確立中における発言権移動の確実性の向上を図り、円滑なコミュニケーションを図ることができる通話端末を提供すること。

【解決手段】多者通話可能な通話端末であって、発言権を有する者の通話端末から送信された音声データを受信して音声認識する受信音声認識手段と、この音声認識結果に基づいて受信した音声データが発言要求であることを検出する発言要求検出手段と、発言要求を検出した場合にその旨をユーザに報知する報知手段と、を備えた。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

多者通話が可能な通話端末であって、
発言権を有する者の通話端末から送信された音声データを受信して音声認識する受信音声認識手段と、

この音声認識結果に基づいて前記受信した音声データが発言要求であることを検出する発言要求検出手段と、

発言要求を検出した場合にその旨をユーザに報知する報知手段と、
を備えたことを特徴とする通話端末。

【請求項 2】

予め設定された発言要求文言を記憶する発言要求文言記憶手段を備えると共に、
前記発言要求検出手段は、前記記憶された発言要求文言に基づいて発言要求を検出する、
ことを特徴とする請求項 1 記載の通話端末。

10

【請求項 3】

前記報知手段は、複数種類の報知を行う、ことを特徴とする請求項 1 又は 2 記載の通話端末。

【請求項 4】

前記報知手段は、前記複数種類の報知を、時間をずらして行う、ことを特徴とする請求項 3 記載の通話端末。

20

【請求項 5】

一定時間内における発言要求検出回数をカウントする発言要求検出回数カウント手段を備えると共に、

前記報知手段は、前記発言要求検出回数に応じて前記複数種類の報知を行う、ことを特徴とする請求項 3 又は 4 記載の通話端末。

【請求項 6】

多者通話が可能な通話端末であって、
他のユーザに対する音声による発言要求の入力を受け付けて音声認識する発言要求音声認識手段と、

この音声認識された音声データに対応する合成音を生成する合成音生成手段と、
この生成された合成音からなる音声データを前記他のユーザの通話端末に送信する発言要求送信手段と、
を備えたことを特徴とする通話端末

30

【請求項 7】

予め設定された発言要求文言を記憶する発言要求文言記憶手段を備えると共に、
前記合成音生成手段は、前記記憶された発言要求文言に基づいて特定の音声データのみに対応する合成音を生成する、
ことを特徴とする請求項 6 記載の通話端末。

【請求項 8】

予め操作者の音声特徴を表す音声特徴データを記憶する音声特徴データ記憶手段を備えると共に、

前記発言要求音声認識手段は、前記音声特徴データに基づいて音声認識を行う、
ことを特徴とする請求項 6 又は 7 記載の通話端末。

40

【請求項 9】

請求項 1 乃至 5 記載の通話端末と、請求項 6 乃至 7 記載の通話端末と、を備えた、ことを特徴とする多者通話システム。

【請求項 10】

多者通話が可能な通話端末にネットワークを介して接続されたサーバコンピュータであって、

前記サーバコンピュータが、前記発言権を有する者の通話端末から送信された他のユー

50

ザに対する音声による発言要求を受信して音声認識する発言要求音声認識手段と、
この音声認識結果に基づいて前記受信した音声データが発言要求であることを検出する
発言要求検出手段と、
発言要求を検出した場合にその旨を前記他のユーザの通話端末に報知する報知制御手段
と、
を備えたことを特徴とするサーバコンピュータ。

【請求項 1 1】

多者通話が可能な通話端末に装備された演算装置に、
発言権を有する者の通話端末から送信された音声データを受信して音声認識する受信音
声認識手段と、
この音声認識結果に基づいて前記受信した音声データが発言要求であることを検出する
発言要求検出手段と、
発言要求を検出した場合にその旨を前記通話端末に装備された報知器を介してユーザに
報知する報知制御手段と、
を実現させるためのプログラム。

10

【請求項 1 2】

多者通話が可能な通話端末に装備された演算装置に、
他のユーザに対する音声による発言要求の入力を受け付けて音声認識する発言要求音声
認識手段と、
この音声認識された音声データに対応する合成音を生成する合成音生成手段と、
この生成された合成音からなる音声データを前記他のユーザの通話端末に送信する発言
要求送信手段と、
を実現させるためのプログラム。

20

【請求項 1 3】

通話端末を用いた多者通話方法であって、
発言要求を受けるユーザの通話端末が、発言権を有する者の通話端末から送信された音
声データを受信して音声認識する受信音声認識工程と、この音声認識結果に基づいて前記
受信した音声データが発言要求であることを検出する発言要求検出工程と、発言要求を検
出した場合にその旨を通信端末に装備された報知器を介してユーザに報知する報知工程と
、
を有することを特徴とする多者通話方法。

30

【請求項 1 4】

前記発言要求検出工程は、発言要求文言記憶手段に予め記憶された発言要求文言に基づ
いて発言要求を検出する、
ことを特徴とする請求項 1 3 記載の多者通話方法。

【請求項 1 5】

前記報知工程は、複数種類の報知を、時間をずらして行う、ことを特徴とする請求項 1
3 又は 1 4 記載の多者通話方法。

【請求項 1 6】

前記報知工程は、一定時間内における発言要求検出回数をカウントすると共に、この発
言要求検出回数に応じて複数種類の報知を行う、
ことを特徴とする請求項 1 3 , 1 4 又は 1 5 記載の多者通話方法。

40

【請求項 1 7】

前記受信音声認識工程の前に、発言権を有する者の通話端末が、他のユーザに対する音
声による発言要求の入力を受け付けて音声認識する発言要求音声認識工程と、この音声認
識された音声データに対応する合成音を生成する合成音生成工程と、この生成された合成
音からなる音声データを前記他のユーザの通話端末に送信する発言要求送信工程と、
を有することを特徴とする請求項 1 3 , 1 4 , 1 5 又は 1 6 記載の多者通話方法。

【請求項 1 8】

前記合成音生成工程は、発言要求文言記憶手段に予め記憶された発言要求文言に基づい

50

て特定の音声データのみに対応する合成音を生成する、
ことを特徴とする請求項 17 記載の多者通話方法。

【請求項 19】

前記発言要求音声認識工程は、音声特徴データ記憶手段に予め記憶された操作者の音声特徴を表す音声特徴データに基づいて音声認識を行う、
ことを特徴とする請求項 17 又は 18 記載の多者通話方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、通話端末にかかり、特に、複数の端末間において多者通話が可能な通話端末に関する。

【背景技術】

【0002】

近年、携帯電話などにおける通信技術の向上により、複数の端末間でグループセッションを確立し、1対N（多数）での通話が可能となっている。このような技術は、例えば、モバイルサービスの標準化団体であるOMA（Open mobile Alliance）にて、Push-to-Talk over Cellular（PoC）という名称で呼ばれている。

【0003】

そして、上述した複数人でのPoC通信は、トランシーバーに近い半2重通話型のコミュニケーション手法にて行われる。つまり、発言権を取得した人間が発言している間、他の人間は傍聴者となる。発言権はPoC通信グループメンバー間で交互にやり取りがなされ、「さんどうぞ」等の口頭での合図（発言要求）により、他の参加者に発言権取得を促すような方法が一般的である。

【0004】

【特許文献1】特開2001-188740号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、上述したような従来の口頭のみによる発言権取得要求では、周囲の雑音で発言要求を聞き取れなかった、または注意散漫で聞いていなかったという場合には、合図（発言要求）が伝わらず円滑なコミュニケーションを阻害する可能性がある。また、PoC通信グループメンバーの中で発言の少ない人間へ発言を促す場合においても、口頭での発言要求だけでなく、同時に、より有効な報知手段にて通知して発言を促すことができれば議論が更に活発化すると考えられる。

【0006】

一方で、上記特許文献1には、電子会議システムにおいて、発言対象者を指定するボタンを押下することで、かかる対象者に発言権が与えられ、当該対象者にその旨を通知する、というシステムが開示されている。しかし、グループセッションにおいて円滑なコミュニケーションを図るためには、従来例と同様に、会話の中で口頭による合図によって他の参加者に発言権取得を促すことが望ましい。

【0007】

このため、本発明は、上記従来例の有する不都合を改善し、特に、グループセッション確立中における発言権移動の確実性の向上を図り、円滑なコミュニケーションを図ることができる通話端末を提供すること、をその目的とする。

【課題を解決するための手段】

【0008】

そこで、本発明の一形態である通話端末は、
多者通話が可能な通話端末であって、

発言権を有する者の通話端末から送信された音声データを受信して音声認識する受信音

10

20

30

40

50

声認識手段と、

この音声認識結果に基づいて受信した音声データが発言要求であることを検出する発言要求検出手段と、

発言要求を検出した場合にその旨をユーザに報知する報知手段と、
を備えたことを特徴としている。

【0009】

上記発明によると、まず、発言権を有する者から発言要求の音声が発せられると、発言要求されたユーザの通話端末にて、音声認識により自動的に発言要求音声を検出される。そして、発言要求された旨が通話端末にて呼び出し音やパイプリータなどによってユーザに報知される。従って、ユーザは、通話内における音声のみでなく、他の方法によっても
10 発言要求があったことを認識することができ、円滑なコミュニケーションを図ることができる。

【0010】

そして、予め設定された発言要求文言を記憶する発言要求文言記憶手段を備えると共に、発言要求検出手段は、記憶された発言要求文言に基づいて発言要求を検出する、ことを特徴としている。これにより、発言要求文言を精度よく検出することができ、より円滑な
コミュニケーションを図ることができる。

【0011】

また、報知手段は、複数種類の報知を行う、ことを特徴としており、特に、報知手段は、複数種類の報知を時間をずらして行う、ことを特徴としている。さらには、一定時間内
20 における発言要求検出回数をカウントする発言要求検出回数カウント手段を備えると共に、報知手段は、発言要求検出回数に応じて複数種類の報知を行う、ことを特徴としている。これにより、通話状況に応じて有効な報知を行うことができると共に、報知がユーザの通話の妨げとなることを抑制でき、さらなる円滑なコミュニケーションを図ることができる。

【0012】

また、本発明における通話端末の他の形態は、

他のユーザに対する音声による発言要求の入力を受け付けて音声認識する発言要求音声認識手段と、

この音声認識された音声データに対応する合成音を生成する合成音生成手段と、
30

この生成された合成音からなる音声データを他のユーザの通話端末に送信する発言要求送信手段と、
を備えたことを特徴としている。

【0013】

上記発明によると、発言権を有する者から発せられた発言要求の音声は、通話端末にて音声認識され、合成音にて他のユーザの通話端末に送信される。従って、聞き取りやすい合成音にて発言要求がなされるため、他の通話端末のユーザは、発言要求を認識しやすくなり、円滑なコミュニケーションを図ることができる。特に、上述したように、発言要求されたユーザの通話端末にて音声認識により自動的に発言要求が検出される場合には、より高精度に発言要求が検出されうる。
40

【0014】

そして、予め設定された発言要求文言を記憶する発言要求文言記憶手段を備えると共に、合成音生成手段は、記憶された発言要求文言に基づいて特定の音声データのみに対応する合成音を生成する、ことを特徴としている。これにより、発言要求文言のみの合成音が生成されるため、端末の処理負担を軽減できる。

【0015】

また、予め操作者の音声特徴を表す音声特徴データを記憶する音声特徴データ記憶手段を備えると共に、発言要求音声認識手段は、音声特徴データに基づいて音声認識を行う、ことを特徴としている。これにより、操作者の癖に応じて発言要求文言を精度よく検出することができ、より円滑なコミュニケーションを図ることができる。
50

【0016】

また、本発明では、上述した発言権を有し発言要求を行う者が所有する通話端末と、この発言要求を受けるユーザの通話端末と、を備えた多者通話システムをも提供している。また、上述した通話端末は、一台の端末にて構成されていることを特徴としている。

【0017】

さらに、本発明の他の形態は、

多者通話が可能で通話端末にネットワークを介して接続されたサーバコンピュータであって、

サーバコンピュータが、発言権を有する者の通話端末から送信された他のユーザに対する音声による発言要求を受信して音声認識する発言要求音声認識手段と、

この音声認識結果に基づいて受信した音声データが発言要求であることを検出する発言要求検出手段と、

発言要求を検出した場合にその旨を他のユーザの通話端末に報知する報知制御手段と、を備えたことを特徴としている。

【0018】

これにより、音声認識処理をサーバコンピュータに実行させることができるため、上述した効果を得ることができると共に、通話端末での処理負担の軽減を図ることができる。

【0019】

また、本発明の他の形態であるプログラムは、

多者通話が可能で通話端末に装備された演算装置に、

発言権を有する者の通話端末から送信された音声データを受信して音声認識する受信音声認識手段と、

この音声認識結果に基づいて受信した音声データが発言要求であることを検出する発言要求検出手段と、

発言要求を検出した場合にその旨を通話端末に装備された報知器を介してユーザに報知する報知制御手段と、を実現させるためのプログラムである。

【0020】

また、本発明であるプログラムの他の形態は、

多者通話が可能で通話端末に装備された演算装置に、

他のユーザに対する音声による発言要求の入力を受け付けて音声認識する発言要求音声認識手段と、

この音声認識された音声データに対応する合成音を生成する合成音生成手段と、

この生成された合成音からなる音声データを他のユーザの通話端末に送信する発言要求送信手段と、

を実現させるためのプログラムである。

【0021】

さらに、本発明では、通話端末を用いた多者通話方法を提供しており、

発言要求を受けるユーザの通話端末が、発言権を有する者の通話端末から送信された音声データを受信して音声認識する受信音声認識工程と、この音声認識結果に基づいて受信した音声データが発言要求であることを検出する発言要求検出工程と、

発言要求を検出した場合にその旨を通信端末に装備された報知器を介してユーザに報知する報知工程と、

を有することを特徴としている。

【0022】

そして、発言要求検出工程は、発言要求文言記憶手段に予め記憶された発言要求文言に基づいて発言要求を検出する、ことを特徴としている。また、報知工程は、複数種類の報知を、時間をずらして行う、ことを特徴としている。さらに、報知工程は、一定時間内における発言要求検出回数をカウントすると共に、発言要求検出回数に応じて複数種類の報知を行う、ことを特徴としている。

【0023】

10

20

30

40

50

さらに、受信音声認識工程の前に、発言権を有する者の通話端末が、他のユーザに対する音声による発言要求の入力を受け付けて音声認識する発言要求音声認識工程と、この音声認識された音声データに対応する合成音を生成する合成音生成工程と、この生成された合成音からなる音声データを他のユーザの通話端末に送信する発言要求送信工程と、を有することを特徴としている。

【0024】

そして、合成音生成工程は、発言要求文言記憶手段に予め記憶された発言要求文言に基づいて特定の音声データのみに対応する合成音を生成する、ことを特徴としている。また、発言要求音声認識工程は、音声特徴データ記憶手段に予め記憶された操作者の音声特徴を表す音声特徴データに基づいて音声認識を行う、ことを特徴としている。

10

【0025】

上記構成の多者通話システム、プログラム、多者通話方法であっても、上述した通話端末と同様に作用するため、上述した本発明の目的を達成することができる。

【発明の効果】

【0026】

本発明は、以上のように構成され機能するので、これによると、他者からの発言要求音声が発音認識により自動的に検出されるため、ユーザは、通話内における音声の発言要求を聞き逃した場合であっても、他の報知方法によって発言要求があったことを認識することができる。従って、多者通話時における円滑なコミュニケーションを図ることができる、という従来にない優れた効果を有する。

20

【発明を実施するための最良の形態】

【0027】

本発明は、多者通話時における円滑なコミュニケーションを図るべく、発言権の移動を促す発言要求である音声データの認識の確実性を向上させる、という点に特徴を有する。以下、実施例では、多者通話が可能で通話端末の一例として携帯電話を挙げ、また、モバイルサービスの標準化団体であるOMA (Open mobile Alliance) にて、Push-to-Talk over Cellular (PoC) という名称で呼ばれている技術を用いて、多者通話を行う場合を一例に挙げて説明する。但し、通話端末は携帯電話に限定されず、また、多者通話を行うシステムは、上記PoCシステムに限定されない。

30

【実施例1】

【0028】

本発明の第1の実施例を、図1乃至図6を参照して説明する。図1は、多者通話システムの全体構成を示す概略図であり、図2は、多者通話を実現する携帯電話の構成を示すブロック図である。図3は、携帯電話に記憶されている情報の一例を示す図である。図4は、多者通話時の様子を示す説明図であり、図5乃至図6は、多者通話時の動作を示すシーケンス図である。

【0029】

[構成]

図1に示すように、本発明における多者通話システムは、複数のユーザA～Dにて操作される携帯電話a～dによって実現される。特に、本実施例では、各携帯電話a～dがネットワークNを介して接続され、上述したPoC通信にて多者通話を行う。なお、本実施例における多者通話は、発言権を有するユーザ(例えば、A)の携帯電話(例えば、a)からの音声データが、他のユーザ(傍聴者)(例えば、B～D)の携帯電話(例えば、b～d)に伝送されることによって実現される。そして、他のユーザが発言するためには、発言権を得る必要があり、当該発言権を有するユーザからの発言要求を受けてから発言する。

40

【0030】

以下では、ユーザA(発言者)が発言権を有することとし、ユーザB(傍聴者)が発言要求を受ける場合を一例に挙げて説明する。つまり、符号aの携帯電話から符号bの携帯電話bに発言要求が送信される場合を説明する。なお、いずれの携帯電話a～dも、同様

50

に構成されており、いずれのユーザ A ~ D であっても、発言権を有する者、あるいは、傍聴者となり得る。

【0031】

図1に携帯電話の構成を示す。ここでは、符号aの携帯電話を一例に挙げて説明するが、他の携帯電話b~cも同様の構成である。この図に示すように、携帯電話aは、装置制御部6を中心に、装置内の各要部へは内部バス13を介して接続されている。装置制御部6には、CPU(中央演算装置)、キャッシュメモリ、割り込みコントローラ、DSP(Digital Signal Processor)、装置全体を制御するためのOSを含み、携帯端末装置全体を制御する役割をもつ。

【0032】

そして、ユーザ(傍聴者)に発言要求を報知する装置(報知器)としては、一般的な携帯電話に装備されている装置内要部の振動駆動装置部1、LED部2、ディスプレイ部3、スピーカ部4が該当する。振動駆動装置部1は、装置全体を振動させる機構を搭載しており、発言要求を受けると振動駆動装置が動作しユーザへ報知する。LED部2は、ユーザが発光を確認可能な装置本体の前面に配置されており、発言要求を受けるとLEDが発光しユーザへ報知する。ディスプレイ部3は、表示制御機能を備えた表示装置である。表示装置は液晶パネルあるいは有機EL(電子蛍光)等の表示装置である。そして、発言要求を受けると、画面全体もしくは一部の表示が変化しユーザへ報知する。このとき表示する画像データは、メモリ5の不揮発領域(ROM)に格納されているユーザデータを使用する。スピーカ部4は、PoC通話中の音声会話を出力する装置であると同時に、発言要求の報知手段にも使用される。発言要求があった場合には、あらかじめ指定された呼び出し音を鳴動させユーザに報知する。呼び出し音としては機械音だけでなく、メモリ5の不揮発領域(ROM)に格納されている音楽・音声データを使用しても良い。

【0033】

なお、上述した報知器である振動駆動装置部1、LED部2、ディスプレイ部3、スピーカ部4の動作は、上述した報知制御手段としての装置制御部6にて制御される。特に、後述するように、傍聴者Bの携帯電話bにて発言要求が検出されたときに、報知器による報知動作がなされるよう制御する。従って、報知器と装置制御部6にて報知手段が構成されている。

【0034】

また、メモリ5には、ROM/RAM共に搭載している。ROMは電話帳、音楽・画像データ等のユーザデータを格納している不揮発性メモリである。また、RAMはCPUがプログラムを実行する上で一時的に必要とされるデータを格納するようになっている。そして、本実施例におけるメモリ5には、予め設定された発言要求キーワード(発言要求文言)が記憶されており、当該メモリ5は、発言要求文言記憶手段として機能している。この発言要求キーワードの一例を、図3に示す。この図に示すように、発言要求キーワードは、一定区切りの単語として格納されており、この各単語を論理演算(AND,OR,NOTなど)により組み合わせることで利用される。例えば、「Aさん、発言してください」という発言要求であれば、図3中のテーブルNo.1とNo.4のAND演算の結果を用いて後述する音声認識処理あるいは発言要求検出処理が行われる。また、「発言の少ない人、発言してください」という要求であれば、タイマ部12の計時機能で発言者毎の発言時間をあらかじめ計測したデータと、図3中のテーブルNo.4とNo.6のAND演算の結果が利用される。そして、これら発言要求キーワードは、発言者Aの携帯電話aにおいては発言要求時に利用され、傍聴者の携帯電話bにおいては、発言要求を検出する際に利用される。

【0035】

また、送信制御部7は、通信用アンテナ9を介して無線によって音声データなどの送信制御を行う回路である。また、受信制御部8は、通信用アンテナ9を介して無線によって音声データなどの受信制御を行う回路である。

【0036】

音声認識部10(発言要求音声認識手段、受信音声認識手段)は、受話部分から入力さ

10

20

30

40

50

れた通話データや、受信制御部 8 を介して他の携帯電話から受信した通話データなど、通話時の音声データに対して、特定キーワードまたは特定制御音の認識・検出を、所定の検出アルゴリズムによって実行する装置またはプログラムにて実現される。音声認識アルゴリズムの詳細については公知のものとして本発明では言及しないが、特に発言権を有するユーザ A にて使用される場合には、その所有者の個性を学習して音声認識精度を向上させることが可能なものとする。つまり、予めユーザの音声特徴を表す音声特徴データを、ROM など（音声特徴データ記憶手段）に記憶しており、これに基づいて音声認識を行う。

【0037】

また、音声認識部 10（発言要求検出手段）は、自分の発言中および他の参加者の発言中は絶えず動作し、会話中に含まれる発言要求となるキーワードを検出する役割をもつ。このとき、上述した ROM に記憶されている図 3 に示すような発言要求キーワードを検出する。そして、傍聴者 B 側にて用いられる場合には、音声認識装置部 10 で発言要求を検出すると、上述したように、装置制御部 6 を介して振動駆動装置部 1、または LED 部 2、またはディスプレイ部 3、またはスピーカ部 4 からユーザへ報知される。

10

【0038】

合成音生成装置部 11（合成音生成手段）は、入力された音声データを加工して出力することが可能な装置である。本発明においては、特に、送信側の携帯電話 a、つまり、PoC 発言者側であるユーザ A の携帯電話 a にて動作し、受信側の電話端末 b（PoC 傍聴者側）の音声認識部 10 の音声認識精度を向上させるためのものである。具体的には、上述したように受話部分から入力された音声データから発言要求キーワードが検出されたときに、そのキーワード自体の合成音を生成して、送信制御部 7 を介して他の通話端末に送信する。つまり、送信制御装置 7 と協働して、発言要求送信手段としても機能する。そして、さらに、合成音生成装置部 11 では、発言要求相手にそれぞれ割り当てられた特定の識別音を付加したり、イントネーションの変更（例えば、地方なまりのイントネーションを標準日本語的なイントネーション化）などの加工を行う。

20

【0039】

また、タイマ部 12 は、時間を計測するための計時装置である。後述するように、発言要求の検出アルゴリズム内で時間計測用途にて使用される。マイク部 14 は、発言者の発言内容を集音する装置であり、受話部分である。

【0040】

[動作]

次に、上記構成の携帯電話の動作を、図 4 乃至図 6 を参照して説明する。以下では、まず、図 4 を参照して全体動作の概要を説明し、続いて、図 5 を参照して送信側であるユーザ A の携帯電話 a の動作を説明し、さらに、図 6 を参照して受信側であるユーザ B の携帯電話 b の動作を説明する。

30

【0041】

<全体動作の概要>

図 4 は、上述した携帯電話 A ~ D を使用し、PoC 会議を行う様子を示したイメージ図である。ここでは、ユーザ A が発言者であり、ユーザ B が傍聴者であって、ユーザ A がユーザ B に発言要求をする様子を説明する。

40

【0042】

図中の発言者（ユーザ A）の携帯電話 a と、傍聴者（ユーザ B）の携帯電話 b は、PoC 通信が確立されており、今、発言者 A が発言しているところである。なお、発言の内容には、傍聴者 B への発言要求となる「 さん、発言してください」、という音声を含んでいるものとする。

【0043】

そして、送信側の携帯電話 a では、発言者 A の肉声が音声データとして入力される（Y1）。この段階の肉声データでは、発言者のしゃべり方の癖、声のトーン、なまり等により、例え同じ発言内容であってもデータ上では多くの差異が存在する。この個人差は、受信側の携帯電話 b の音声認識精度の低下を招く要因であり、認識精度が低い場合には予期

50

せぬ状況で報知動作が誤作動してしまう問題がある。従って、本実施例では、上述したように、発言者 A の声の特徴を学習済みの送信側携帯電話 a の音声認識部を用いて、肉声データ中の発言要求部分の検出を行い、発言要求部分の音声データを合成音に変換する (Y 2)。すると、合成音への変換によって肉声データで介在していた個体差は除去され、受信側の携帯電話 b での認識精度を向上させる効果が期待できる。なお、合成音生成手法としては、音声データに特定の識別音を付加するということでもよく、発言のイントネーション変更 (なまりなどの発言の癖を除去) であってもよい。

【 0 0 4 4 】

このように、音声データは発言要求部分が合成音化され、PoC通信網を伝わって受信側端末装置へと伝送される (Y 3)。すると、受信側の携帯電話 b 内部では、受信した合成音による音声データ (「 さん、発言してください 」) をスピーカ等の鳴動装置により外部出力する (Y 4) と共に、音声認識部が会話中の発言要求検出を絶えず行う。そして、合成音化された発言要求を音声認識部が検出した場合 (Y 5) には、発言要求の報知動作を行う (Y 6)。なお、符号 Y 4 の会話出力動作は省略し、符号 Y 6 の報知動作のみを行ってもよい。

【 0 0 4 5 】

< 送信側携帯電話の動作 >

次に、各携帯電話 a , b における動作について詳述する。まず、図 5 のシーケンス図を参照して、発言者 A の携帯電話 a の動作を詳述する。

【 0 0 4 6 】

発言者 A による発言は、マイク部より入力され音声データへと変換され (ステップ S 1)、メモリ部へと逐次転送される (ステップ S 2)。音声データは、メモリ部 5 の音声認識用のバッファに格納され (ステップ S 3)、音声認識部 1 0 ではそのバッファに格納された音声データから音声認識処理を行う (ステップ S 4)。音声認識処理は会話が行われている間は常に動作し、発言要求となる特定キーワードの検出を行う (ステップ S 5)。

【 0 0 4 7 】

そして、特定キーワードの検出がされた場合は、その音声部分のデータを合成音生成部 1 1 へと転送し (ステップ S 7)、合成音生成部 1 1 にて合成音へと変換した後 (ステップ S 8)、送信制御部 7 へと転送される (ステップ S 9)。一方、特定キーワードが検出されていない音声データは、そのままの肉声データとして送信制御部 9 へと転送される (ステップ S 6)。送信制御部 7 では、転送されてきたデータを符号化し、通信用アンテナ 9 からデータを傍聴者 B の携帯電話 b に送信する (ステップ S 1 0)。

【 0 0 4 8 】

< 受信側携帯電話の動作 >

次に、図 6 のシーケンス図を参照して、傍聴者 B の携帯電話 b の動作を詳述する。まず、装置制御部 6 では、常時、発言要求の検出待ち状態にある (ステップ S 2 0)。そして、通信用アンテナ 9 にて、上述した発言者 A の携帯電話 a からデータを受信すると、受信制御部 8 は音声データを復号化する (ステップ S 2 1)。その後、受信された音声データは逐次転送され (ステップ S 2 2)、スピーカ部 4 で会話出力される (ステップ S 2 9 , S 3 0) と共に、メモリ部 5 の音声認識用バッファへ格納される (ステップ S 2 3)。音声認識部 1 0 では、そのバッファに格納された音声データから音声認識処理を行う (ステップ S 2 4)。音声認識処理は会話が行われている間は常に動作し、発言要求となる特定キーワードの検出を行う (ステップ S 2 5)。

【 0 0 4 9 】

そして、発言要求の検出がされた場合は、直ちに装置制御部 6 へと特定キーワードヒットの通知が行われ (ステップ S 2 6)、装置制御部 6 はその通知を受け、スピーカ部 4、ディスプレイ部 3、LED部 2、振動制御部 1 へと報知動作要求を行う (ステップ S 2 7)。すると、この報知動作要求に応じて、各部 1 ~ 4 が報知作動する (ステップ S 2 8)。なお、発言要求の報知動作は、各部で同時に行ってもよく、あるいは、単独に時間をずらして行ってもよい。

10

20

30

40

50

【0050】

以上のように、本実施例では、PoC会議中の「さんどうぞ」等の音声による発言要求を、受信側の携帯電話bで音声認識により自動的に検出し、装置の振動、LCD表示、LED点灯、制御音呼応等の通知手段を動作させることが可能となる。つまり、ユーザは音声だけでなく、端末装置の他動作によって発言要求を知る機会を得ることができる。従って、従来よりも確実に相手に発言要求を通知させることにより、周囲の雑音で発言要求を聞き取れなかった、または注意散漫で発言要求を聞いていなかったという状況を解消させ、円滑なコミュニケーションを図ることができる。

【0051】

さらに、本実施例では、送信側の電話端末aでも音声認識を行い、認識した発言要求部分の音声データを合成音へと変換することで発言者のしゃべり方の癖、なまり等の個人差を一般化させることができる。これにより、受信側の携帯電話bによる音声認識精度が向上し、さらなるコミュニケーションの円滑化を図ることができる。

【0052】

なお、上記とは異なり、送信側の携帯電話aからの発言要求音声は、合成音化されずに肉声のまま受信側の携帯電話bに送信されてもよい。この場合には、受信側の携帯電話bでは、肉声に対して音声認識を行い、発言要求の検出が行われる。このようにしても、受信側の携帯電話bにて、音声認識により自動的に発言要求音声を検出されるため、円滑なコミュニケーションを図ることができる。

【0053】

また、上記とは異なり、受信側の携帯電話bでは音声認識が行われず、送信側の携帯電話aからの発言要求音声は合成音化されて受信側の携帯電話bに送信されるのみであってもよい。このようにしても、受信側の携帯電話bに対して、傍聴者が聞き取りやすい（認識しやすい）合成音にて発言要求がなされるため、円滑なコミュニケーションを図ることができる。

【0054】

なお、上述した構成は、携帯電話の構成を変更するのみで実現可能であり、現状の多者通話を可能とするPoC通信システムの改良を必要としないため、コスト面での負担が軽減される。

【実施例2】

【0055】

次に、本発明の第2の実施例を、図7を参照して説明する。図7は、受信側の携帯電話bの報知動作を示すシーケンス図である。

【0056】

仮に、音声認識の誤認識により報知動作が行われてしてしまう場合には、かえってコミュニケーションを阻害しかねない。とりわけ、振動駆動装置が動作してしまっただけではユーザの不満は大きい。このため、誤報によるユーザへの影響を低減させるための制御手法として、本実施例では、ユーザ報知の強さが弱い順に時間を置いて段階的に動作させることとする。例えば、上述したように、振動、LCD表示、LED点灯、音の鳴動という4つの報知手段があったとすると、LED点灯、LCD表示、音の鳴動、振動の順に行うように制御を行う。

【0057】

具体的な動作を、図7を参照して説明する。ここでは、受信側の携帯電話bが、送信側の携帯電話aからの発言要求の検出を待ち状態にあり（ステップS40）、装置制御部6が発言要求を受け取り、報知手段動作部1、および報知手段動作部2を制御させようとしている状況にある。そして、装置制御部6が発言要求を受け取ると（ステップS41）、タイマ部12へタイマ開始要求を行い（ステップS42）、タイマ部12は一定時間のタイマ計測を行う（ステップS43）。タイマ計測を終えると、タイマ部12は装置制御部6にカウントアップ完了通知を行う（ステップS44）。そこではじめて装置制御部6は、報知動作の要求を報知手段動作部1へ通知する（ステップS45）し、報知動作が行われる（ステップS46）。このとき、1回目は、LED部2による発光といった傍聴者Bに

10

20

30

40

50

対する報知強さが弱い報知器を用いて行われる。

【0058】

同様にして、さらに時間が計測され、一定時間が経過すると、ステップS47～ステップS51に示すように、2回目の報知が行われる。このときは、スピーカ部4から音による報知を行うといった、さらに報知度が強い報知器を用いて行われる。

【0059】

このように、複数種類の報知を、時間をずらして行い、さらには、ユーザに対する報知強さを変えて段階的に報知することで、音声認識の誤認識により報知動作が行われてしてしまう場合にフェールセーフ的に作動し、誤作動によるユーザの不満を低減させる効果がある。

【実施例3】

【0060】

次に、本発明の第3の実施例を、図8を参照して説明する。図8は、受信側の携帯電話bの報知動作を示すシーケンス図である。

【0061】

本実施例では、一定時間内における発言要求の検出回数に応じて、報知方法を変えて報知する、という点に特徴を有する。つまり、装置制御部6は、一定時間内における発言要求(その一部のキーワード)が検出された回数をカウントする機能(発言要求検出回数カウント手段)を有すると共に、その回数に応じて各種の報知を行う、という機能を有する。これは、発言要求が短時間に繰り返し行われるという傾向を考慮してのことである。例えば、傍聴者Bが「さん」であった場合に、1回目の発言要求は「さんコメントを頂きたいと思います。」、2回目の発言要求は「では、さんお願いします。」、3回目の発言要求は「さん、聞いてますか?」といったような会話を想定しており、この場合には、検出回数をカウントする特定キーワードは、「さん」となる。

【0062】

図8に示すように、まず、装置制御部6では、1回目の検出待ち(ヒット待ち)状態となっている(ステップS60)。そして、特定キーワードの検出の通知がなされたら(ステップS61)、タイマ部12にタイマ開始要求がなされ(ステップS62)、タイマ部12ではタイマ計測が開始され(ステップS63)、上記特定キーワードのカウントが開始される。同時に、装置制御部6は2回目のヒット待ち状態となる(ステップS64)。その後、装置制御部6が2回目のヒット待ち状態中に、特定キーワードヒットの通知(2回目)がなされたら(ステップS65)、まずは比較的弱い報知手段を行う報知手段動作部1へ報知動作の要求(ステップS66)を通知し、報知手段を動作させる(ステップS67)。さらに、もう一度、特定キーワードヒットの通知(3回目)がなされたら(ステップS68)、強い報知手段を行う報知手段動作部2へ報知動作の要求(ステップS69)を通知し、報知手段を動作させる(ステップS70)。

【0063】

なお、図8の点線内に示すように、2回目やそれ以降のキーワードヒットが通知される前にタイマカウントアップが完了した場合には、タイマ部12は装置制御部6へカウントアップ完了通知を行い(ステップS71)、装置制御部6は1回目のヒット待ち状態へと遷移する(ステップS72)。

【0064】

これにより、より確実に発言要求を通知することができ、さらなる円滑なコミュニケーションを図ることができる。

【実施例4】

【0065】

次に、本発明の第3の実施例を、図9を参照して説明する。上記では、発言要求の音声認識処理を、各携帯電話a, bにて行っていたが、かかる処理をPoCサーバシステム20で実行させてもよい。つまり、上述した携帯電話bが有する音声認識部10(発言要求音声認識手段、発言要求検出手段)を、サーバシステム20が備えており、さらに、発言要

10

20

30

40

50

求を検出した旨を受信側の携帯電話 b に通知する機能（報知制御手段）を備えている。

【0066】

その動作を、図9を参照して説明する。まず、発言者（A）から傍聴者（B）へ「さん、発言してください」という発言要求を入力すると、送信側の携帯電話 a では、入力された発言（Y11）をそのままPoCサーバシステム20に伝送する（Y12, Y13）。PoCサーバシステム20では、音声認識部を備えており、上記実施例にて説明した電話端末 b と同様に発言要求を検出し（Y14）、要求対象者となる者の携帯端末装置へと発言要求を伝える制御信号を送信する（Y15）。このとき、音声データもそのまま送信する。そして、この制御信号を受け取った携帯端末 b では、会話出力中（Y16）に発言要求の報知（Y17, Y18）が動作する。

10

【0067】

これにより、音声認識処理をサーバシステム20に実行させることができるため、上述した効果を得ることができると共に、携帯電話 a, b での処理負担の軽減を図ることができる。

【実施例5】

【0068】

次に、本発明の第5の実施例を、図10乃至図11を参照して説明する。図10乃至図11は、携帯電話の構成を示す図である。

【0069】

本実施例における携帯電話は、特に、発言権を有し、他のユーザに対して発言要求を行う送信側の携帯電話 a である。そして、基本的には、図2に示すように上述した実施例における携帯電話 a, b と同様の構成であるが、図10及び図11に示すように、さらに、キーボード部15を装備している点で異なる。そして、このキーボード部15の各キーには、図11に示すように、PoC会議参加者と、参加者を判別可能な制御音声と、が関連付けられており、かかる関連付け情報があらかじめ制御音変換テーブル51として、メモリ部5に格納されている。なお、かかる関連付けは、ユーザAによって行われる。

20

【0070】

そして、発言者Aは、発言要求を行いたい相手がいた場合、その者に対応したキーを押下（Y21）することで、対応制御音が合成音生成部11にて生成される。そして、かかる制御音は、各参加者で個別のものであるため、受信側の電話端末 b では、音声認識処理により、他のユーザBに対する発言要求であることを検出することができる。つまり、制御音は、上述した発言要求の合成音として機能する。

30

【0071】

このようにすることで、送信側の携帯電話 a では、「さんお願いします」という発言要求を喋らずとも、キー押下のみで発言要求を行うのと同じ効果を得られる。なお、この機能は、上述した音声による発言要求を発する際に、補助的に利用されてもよい。つまり、音声に基づいて生成された合成音による発言要求と共に、上記制御音を送信してもよい。

【産業上の利用可能性】

【0072】

本発明は、携帯電話機、PHS（Personal Handyphone System）、PDA（Personal Data Assistance, Personal Digital Assistants: 個人向け携帯型情報通信機器）等の携帯端末装置やPC（Personal Computer）等の通信機能を備えており、多者通話が可能である端末装置に利用することができ、産業上の利用可能性を有する。

40

【図面の簡単な説明】

【0073】

【図1】多者通話システムの全体構成を示す概略図である。

【図2】実施例1における携帯電話の構成を示すブロック図である。

【図3】携帯電話に記憶されている発言要求キーワードの一例を示す図である。

【図4】実施例1における多者通話時の様子を示す説明図である。

50

【図5】実施例1における多者通話時の動作を示すシーケンス図である。

【図6】実施例1における多者通話時の動作を示すシーケンス図である。

【図7】実施例2における多者通話時の動作を示すシーケンス図である。

【図8】実施例3における多者通話時の動作を示すシーケンス図である。

【図9】実施例4における多者通話時の様子を示す説明図である。

【図10】実施例5における携帯電話の構成を示すブロック図である。

【図11】実施例5における多者通話時の様子を示す説明図である。

【符号の説明】

【0074】

1 振動駆動装置部

2 LED部

3 ディスプレイ部

4 スピーカ部

5 メモリ

6 装置制御部

7 送信制御部

8 受信制御部

9 通信用アンテナ

10 音声認識部

11 合成音生成装置部

12 タイマ部

13 内部バス

14 マイク部

15 キーボード部

A ユーザ（発言者）

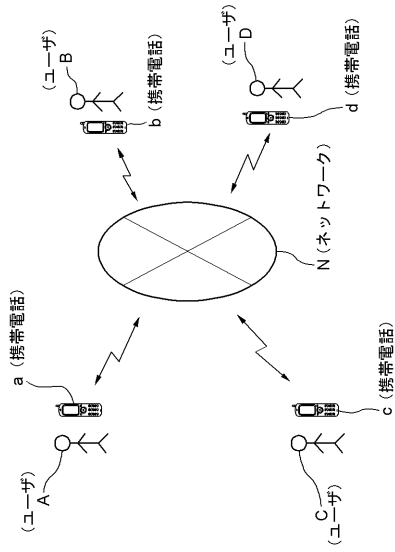
B ユーザ（傍聴者）

a, b, c, d 携帯電話

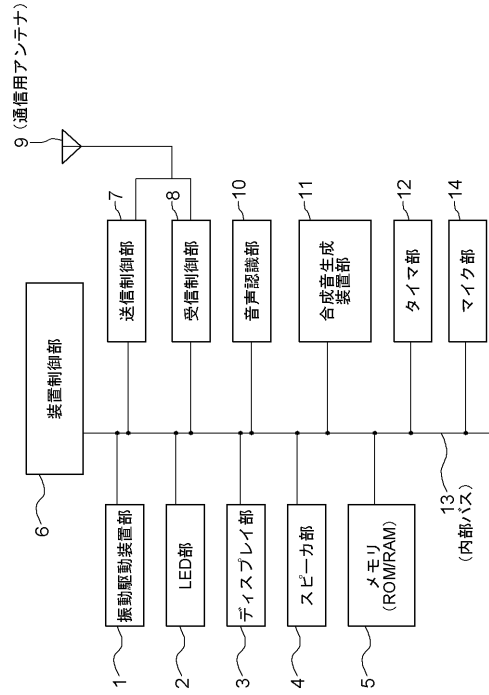
10

20

【 図 1 】



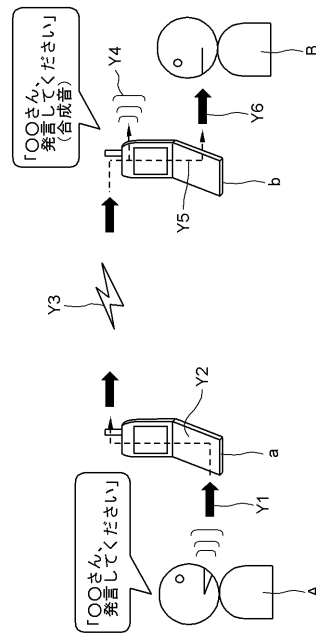
【 図 2 】



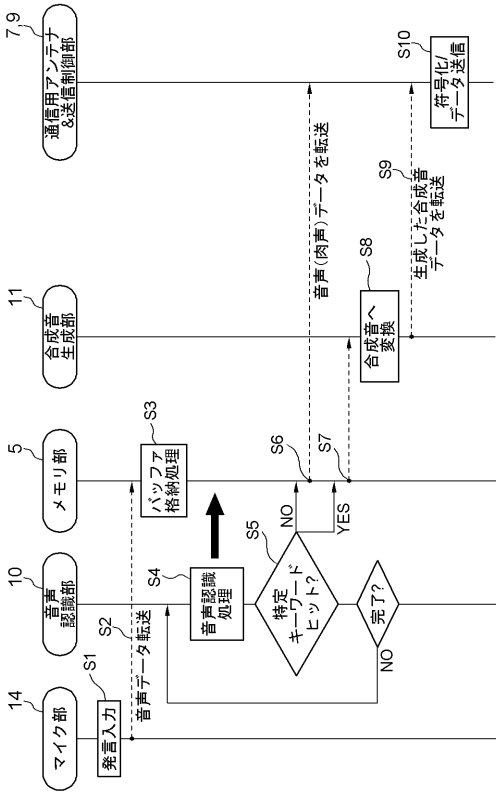
【 図 3 】

ワードNo	発言要求キーワード
1	「Aさん」
2	「Bさん」
3	「Cさん」
4	「発言して下さい」
5	「みなさん」
6	「発言の少ない」
7	「発言の多い」
8	「誰か」
...	...
...	...
...	...

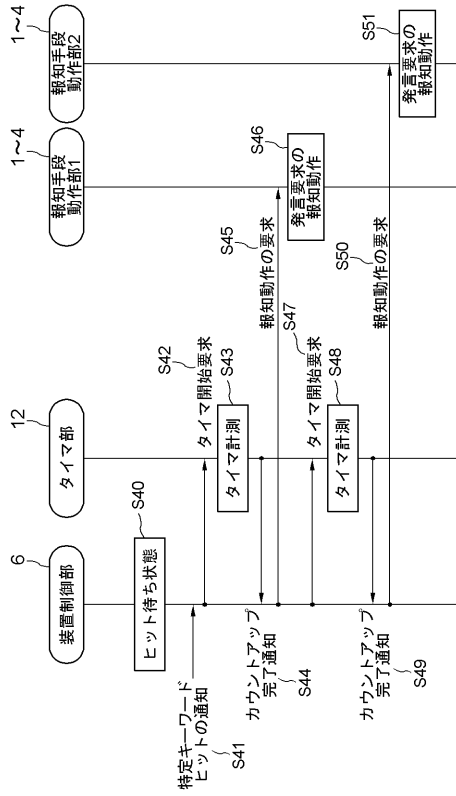
【 図 4 】



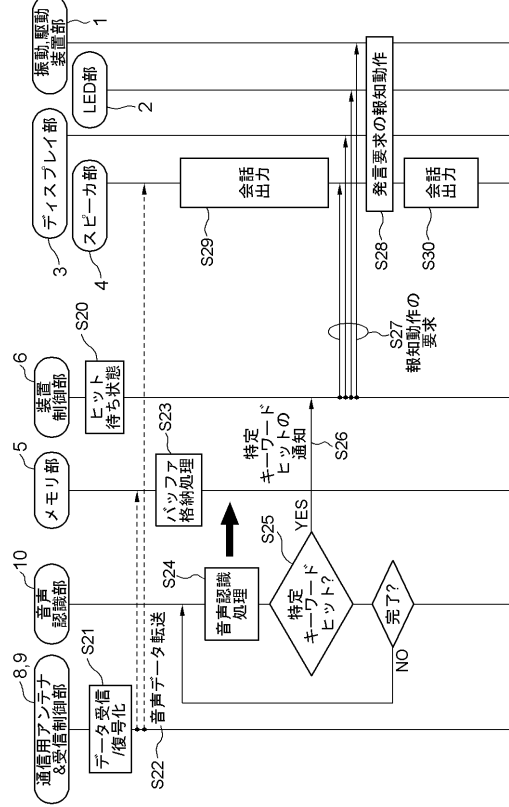
【図 5】



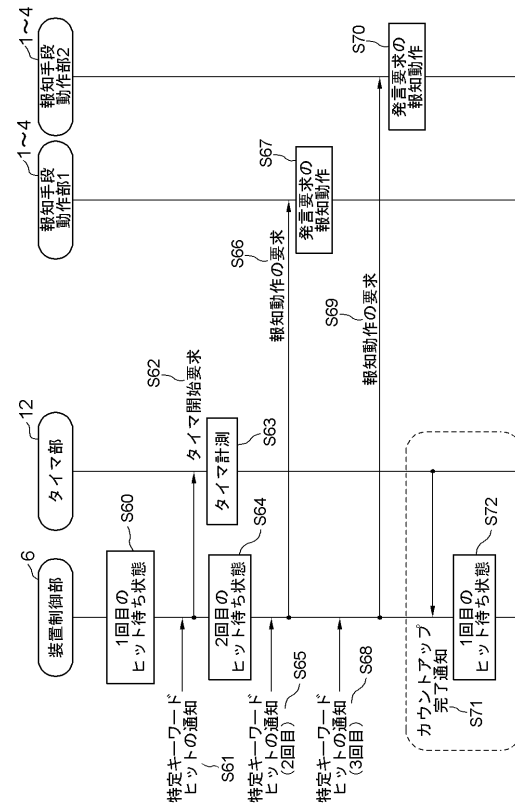
【図 7】



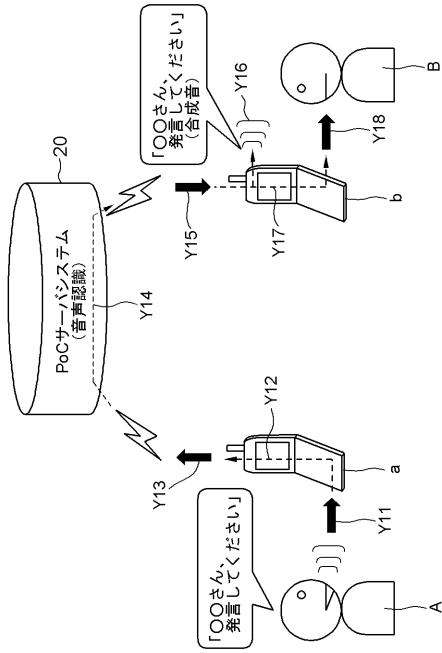
【図 6】



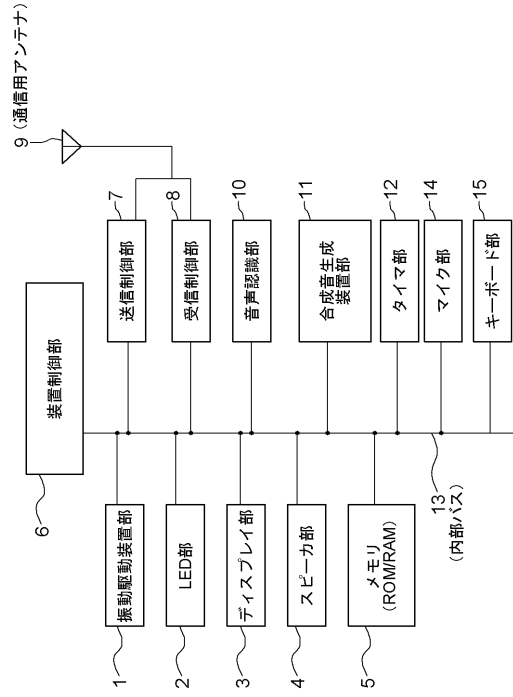
【図 8】



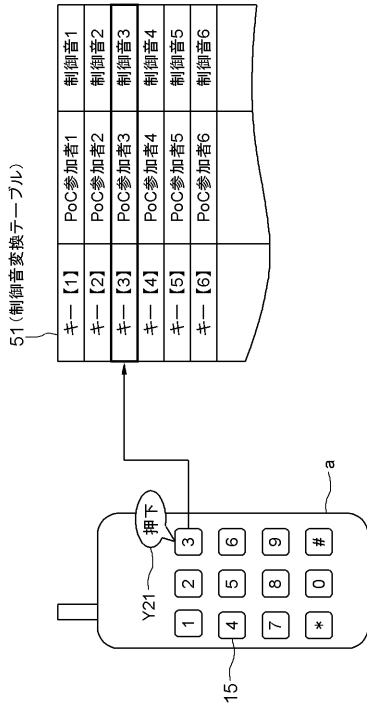
【 図 9 】



【 図 10 】



【 図 11 】



フロントページの続き

(51) Int.Cl.

F I

テーマコード(参考)

G 1 0 L 13/00 1 0 0 C

G 1 0 L 15/00 2 0 0 U