

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6573676号
(P6573676)

(45) 発行日 令和1年9月11日(2019.9.11)

(24) 登録日 令和1年8月23日(2019.8.23)

(51) Int. Cl.	F I		
HO4M 3/56 (2006.01)	HO4M 3/56	Z	
HO4M 3/00 (2006.01)	HO4M 3/00	B	
HO4M 3/42 (2006.01)	HO4M 3/42	P	

請求項の数 17 (全 31 頁)

(21) 出願番号	特願2017-540583 (P2017-540583)	(73) 特許権者	503433420
(86) (22) 出願日	平成27年1月30日 (2015.1.30)		華為技術有限公司
(65) 公表番号	特表2018-509056 (P2018-509056A)		HUAWEI TECHNOLOGIES CO., LTD.
(43) 公表日	平成30年3月29日 (2018.3.29)		中華人民共和国 518129 広東省深 ▲チェン▼市龍崗区坂田 華為総部▲ベン ▼公楼
(86) 国際出願番号	PCT/CN2015/071966		Huawei Administrati on Building, Bantia n, Longgang Distric t, Shenzhen, Guangd ong 518129, P. R. Ch ina
(87) 国際公開番号	W02016/119226	(74) 代理人	100110364
(87) 国際公開日	平成28年8月4日 (2016.8.4)		弁理士 実広 信哉
審査請求日	平成29年9月8日 (2017.9.8)		最終頁に続く

(54) 【発明の名称】 多者通話において音声テキストに変換するための方法および装置

(57) 【特許請求の範囲】

【請求項1】

サーバに適用される、多者通話において音声テキストに変換するための方法であって、

少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するステップであって、前記音声 - テキスト変換要求が、第1の識別子および第2の識別子を含む、ステップと、

前記少なくとも2つの端末によって送信された前記音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、前記少なくとも2つの端末にセッションを割り当てるステップと、

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するステップであって、前記多者通話が、1回のセッションに対応する、ステップと、

前記音声ストリームをテキストに変換するステップと、

前記多者通話において端末に前記テキストを送信するステップと

を含む方法。

【請求項2】

前記少なくとも2つの端末にセッションを割り当てる前記ステップが、

データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、前記第1の端末によって送信された前記音声 - テキスト変

換要求の前記第1の識別子または前記第2の識別子に対応する、前記データベース内のセッションを前記第1の端末に割り当てるステップであって、前記第1の端末が、前記少なくとも2つの端末のうちのいずれかの端末であり、前記データベースが、少なくとも1つのセッションおよび該少なくとも1つのセッションに対応する端末識別子を含む、ステップと、

前記データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、前記第2の端末によって送信された前記音声 - テキスト変換要求の前記第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、かつ前記第2の端末によって送信された前記音声 - テキスト変換要求の前記第2の識別子が、前記第3の端末によって送信された前記音声 - テキスト変換要求の第2の識別子と同じである場合、前記第2の端末および前記第3の端末に同じ新しいセッションを割り当てるステップと
10
を含む、請求項1に記載の方法。

【請求項3】

前記少なくとも2つの端末にセッションを割り当てる前記ステップの後に、
割り当てられたセッションが新しいセッションである場合、前記割り当てられたセッションと前記少なくとも2つの端末の識別子との対応関係を確立し、前記対応関係をデータベースに追加するステップ、または
前記割り当てられたセッションが前記データベース内のセッションである場合、追加された識別子が前記割り当てられたセッションに対応するように、前記少なくとも2つの端末の前記識別子のうちにあり、かつ前記データベースに含まれない端末の識別子を前記データベースに追加するステップ
20
をさらに含む、請求項1または2に記載の方法。

【請求項4】

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信する前記ステップの後に、
第5の端末によって送信された終了メッセージを受信するステップと、
データベースから前記第5の端末の識別子を削除するステップと、
セッションに対応する端末の識別子が前記データベースにおいて空である場合、前記データベースから前記セッションを削除するステップと
30
をさらに含む、請求項1から3のいずれか一項に記載の方法。

【請求項5】

前記テキストが、前記少なくとも1つの端末の識別子を含むか、または、前記テキストが、前記少なくとも1つの端末の識別子およびユーザ名を含む、請求項1から4のいずれか一項に記載の方法。

【請求項6】

多者通話において音声をテキストに変換するための装置であって、
少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するように構成された第1の受信ユニットであって、前記音声 - テキスト変換要求が、第1の識別子および第2の識別子を含む、第1の受信ユニットと、
40
前記少なくとも2つの端末によって送信された前記音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、前記少なくとも2つの端末にセッションを割り当てるように構成された割り当てユニットと、
1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するように構成された第2の受信ユニットであって、前記多者通話が、1回のセッションに対応する、第2の受信ユニットと、

前記音声ストリームをテキストに変換するように構成された変換ユニットと、
前記多者通話において端末に前記テキストを送信するように構成された送信ユニットと
50
を備える装置。

【請求項 7】

前記割当ユニットが、

データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、前記第1の端末によって送信された前記音声 - テキスト変換要求の前記第1の識別子または前記第2の識別子に対応する、前記データベース内のセッションを前記第1の端末に割り当てることであって、前記第1の端末が、前記少なくとも2つの端末のうちのいずれかの端末であり、前記データベースが、少なくとも1つのセッションおよび該少なくとも1つのセッションに対応する端末識別子を含む、ことと、

前記データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、前記第2の端末によって送信された前記音声 - テキスト変換要求の前記第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、かつ前記第2の端末によって送信された前記音声 - テキスト変換要求の前記第2の識別子が、前記第3の端末によって送信された前記音声 - テキスト変換要求の第2の識別子と同じである場合、前記第2の端末および前記第3の端末に同じ新しいセッションを割り当てることと

を行うようにさらに構成される、請求項6に記載の装置。

【請求項 8】

割り当てられたセッションが新しいセッションである場合、前記割り当てられたセッションと前記少なくとも2つの端末の識別子との対応関係を確立し、前記対応関係をデータベースに追加するように構成された追加ユニットであって、

前記割り当てられたセッションが前記データベース内のセッションである場合、追加された識別子が前記割り当てられたセッションに対応するように、前記少なくとも2つの端末の前記識別子のうちにあり、かつ前記データベースに含まれない端末の識別子を前記データベースに追加するようにさらに構成された追加ユニット

をさらに備える、請求項6または7に記載の装置。

【請求項 9】

第5の端末によって送信された終了メッセージを受信するように構成された第3の受信ユニットと、

データベースから前記第5の端末の識別子を削除するように構成された削除ユニットであって、

セッションに対応する端末の識別子が前記データベースにおいて空である場合、前記データベースから前記セッションを削除するようにさらに構成された削除ユニットと

をさらに備える、請求項6から8のいずれか一項に記載の装置。

【請求項 10】

前記テキストが、前記少なくとも1つの端末の識別子を含むか、または、前記テキストが、前記少なくとも1つの端末の識別子およびユーザ名を含む、請求項6から9のいずれか一項に記載の装置。

【請求項 11】

多者通話において音声を変換するための装置であって、

少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するように構成された第1の受信器であって、前記音声 - テキスト変換要求が、第1の識別子および第2の識別子を含む、第1の受信器と、

前記少なくとも2つの端末によって送信された前記音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、前記少なくとも2つの端末にセッションを割り当てるように構成されたプロセッサと、

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するように構成された第2の受信器であって、前記多者通話が、1回のセッションに対応する、第2の受信器と

を備え、

10

20

30

40

50

前記プロセッサが、前記音声ストリームをテキストに変換するようにさらに構成され、前記装置が、

前記多者通話において端末に前記テキストを送信するように構成された送信器をさらに備える、装置。

【請求項 1 2】

前記プロセッサが、

データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、前記第1の端末によって送信された前記音声 - テキスト変換要求の前記第1の識別子または前記第2の識別子に対応する、前記データベース内のセッションを前記第1の端末に割り当てることであって、前記第1の端末が、前記少なくとも2つの端末のうちのいずれかの端末であり、前記データベースが、少なくとも1つのセッションおよび該少なくとも1つのセッションに対応する端末識別子を含む、ことと、

前記データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、前記第2の端末によって送信された前記音声 - テキスト変換要求の前記第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、かつ前記第2の端末によって送信された前記音声 - テキスト変換要求の前記第2の識別子が、前記第3の端末によって送信された前記音声 - テキスト変換要求の第2の識別子と同じである場合、前記第2の端末および前記第3の端末に同じ新しいセッションを割り当てることと

を行うようにさらに構成される、請求項11に記載の装置。

【請求項 1 3】

前記プロセッサが、

割り当てられたセッションが新しいセッションである場合、前記割り当てられたセッションと前記少なくとも2つの端末の識別子との対応関係を確立し、前記対応関係をデータベースに追加することと、

前記割り当てられたセッションが前記データベース内のセッションである場合、追加された識別子が前記割り当てられたセッションに対応するように、前記少なくとも2つの端末の前記識別子のうちにあり、かつ前記データベースに含まれない端末の識別子を前記データベースに追加することと

を行うようにさらに構成される、請求項11または12に記載の装置。

【請求項 1 4】

第5の端末によって送信された終了メッセージを受信するように構成された第3の受信器をさらに備え、

前記プロセッサが、データベースから前記第5の端末の識別子を削除するようにさらに構成され、

前記プロセッサが、セッションに対応する端末の識別子が前記データベースにおいて空である場合、前記データベースから前記セッションを削除するようにさらに構成される、請求項11から13のいずれか一項に記載の装置。

【請求項 1 5】

前記テキストが、前記少なくとも1つの端末の識別子を含むか、または、前記テキストが、前記少なくとも1つの端末の識別子およびユーザ名を含む、請求項11から14のいずれか一項に記載の装置。

【請求項 1 6】

請求項1から5のいずれか一項に記載の方法をコンピュータに実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 1 7】

請求項1から5のいずれか一項に記載の方法をコンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本発明は、通信技術の分野に関し、特に、多者通話において音声を変換するための方法および装置に関する。

【背景技術】

【0002】

通信技術の継続的な発展に伴い、電話は、日常生活において不可欠となっており、電話は、より多くの機能を有するようになってきている。電話は、手っ取り早く、便利なため、多くの重要なビジネス交渉、仕事の手配、およびミーティングなどは電話で行われている。電話によって前述の業務を行う信頼性を確保するために、会話の終了後に、通話記録（すなわち、通話テキスト）を取得する必要がある。

【0003】

現在のところ、通話テキストを取得する主な方法は、以下のようなもの、すなわち、手動録音を最初に実行した後、手動で録音を聞き取り、録音をテキストにまとめるというものであり、このようにして取得された通話テキストの精度は、比較的高くはあるが、この方法は、時間と労力がかかるものである。あるいは、通話処理において、通話の音声ストリームをサンプリングし、サンプリングによって取得された音声ストリームを音声認識エンジンに送信するというものもある。音声認識エンジンは、音声情報をテキスト情報に変換し、変換されたテキストを、ユーザが使用する端末に送信する。この処理は、手動操作を必要とせず、人間の労力と時間を節約する。しかしながら、回線交換ドメインを用いて取得される音声情報のサンプリングレートは、8KHzであるため、音声認識エンジンによる音声認識の効果は低く、音声認識エンジンで変換されるテキストの精度は低い。

【0004】

結論として、現時点では、多者通話において音声から変換されたテキストを、時間を節約する容易な方法で正確に取得する方法は存在していない。

【発明の概要】

【課題を解決するための手段】

【0005】

本発明の実施形態は、多者通話において音声を変換するための方法および装置を提供する。これは、多者通話において音声から変換されたテキストを、時間を節約する容易な方法で正確に取得することができる。

【0006】

前述の目的を達成するために、本発明の実施形態では、以下の技術的解決策が採用される。

【0007】

第1の態様によれば、サーバに適用される、多者通話において音声を変換するための方法が提供され、本方法は、

少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するステップであって、音声 - テキスト変換要求が、第1の識別子および第2の識別子を含む、ステップと、

少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てるステップと、

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するステップであって、多者通話が、1回のセッションに対応する、ステップと、

音声ストリームをテキストに変換するステップと、

多者通話において端末にテキストを送信するステップと

を含む。

【0008】

第1の態様に関連して、第1の態様の第1の可能な実施形態では、少なくとも2つの端末にセッションを割り当てるステップは、

データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子に対応する、データベース内のセッションを第1の端末に割り当てるステップであって、第1の端末が、少なくとも2つの端末のうちのいずれかの端末であり、データベースが、少なくとも1つのセッションおよび該少なくとも1つのセッションに対応する端末識別子を含む、ステップと、

データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、かつ第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じである場合、第2の端末および第3の端末に同じ新しいセッションを割り当てるステップとを含む。

【0009】

第1の態様または第1の態様の第1の可能な実施態様に関連して、第1の態様の第2の可能な実施態様では、テキストは、少なくとも1つの端末の識別子を含むか、または、テキストは、少なくとも1つの端末の識別子およびユーザ名を含む。

【0010】

第1の態様または第1の態様の第1の可能な実施態様または第1の態様の第2の可能な実施態様に関連して、第1の態様の第3の可能な実施態様では、少なくとも2つの端末にセッションを割り当てるステップの後に、本方法は、

割り当てられたセッションが新しいセッションである場合、割り当てられたセッションと少なくとも2つの端末の識別子との対応関係を確認し、対応関係をデータベースに追加するステップ、または

割り当てられたセッションがデータベース内のセッションである場合、追加された識別子が割り当てられたセッションに対応するように、少なくとも2つの端末の識別子のうちにあり、かつデータベースに含まれない端末の識別子をデータベースに追加するステップをさらに含む。

【0011】

第1の態様または第1の態様の第1から第3の可能な実施態様のいずれかに関連して、第1の態様の第4の可能な実施態様では、1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するステップの後に、本方法は、

第5の端末によって送信された終了メッセージを受信するステップと、

データベースから第5の端末の識別子を削除するステップと、

セッションに対応する端末の識別子がデータベースにおいて空である場合、データベースからセッションを削除するステップと

をさらに含む。

【0012】

第2の態様によれば、多者通話において音声をテキストに変換するための装置が提供され、本装置は、

少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するように構成された第1の受信ユニットであって、音声 - テキスト変換要求が、第1の識別子および第2の識別子を含む、第1の受信ユニットと、

少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てるように構成された割当ユニットと、

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するように構成された第2の受信ユニットであって、多者通話が、1回のセッションに対応する、第2の

10

20

30

40

50

受信ユニットと、

音声ストリームをテキストに変換するように構成された変換ユニットと、
多者通話において端末にテキストを送信するように構成された送信ユニットと
を含む。

【 0 0 1 3 】

第2の態様に関連して、第2の態様の第1の可能な実施態様では、割り当ユニットは、

データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子
または第2の識別子を含む場合、第1の端末によって送信された音声 - テキスト変換要求の
第1の識別子または第2の識別子に対応する、データベース内のセッションを第1の端末に
割り当てることであって、第1の端末が、少なくとも2つの端末のうちのいずれかの端末で
あり、データベースが、少なくとも1つのセッションおよび該少なくとも1つのセッション
に対応する端末識別子を含む、ことと、

10

データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子
および第2の識別子を含まず、第2の端末によって送信された音声 - テキスト変換要求の第
1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同
じであり、かつ第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、
第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じである場合、
第2の端末および第3の端末に同じ新しいセッションを割り当てることと

を行うように構成される。

【 0 0 1 4 】

20

第2の態様または第2の態様の第1の可能な実施態様に関連して、第2の態様の第2の可能
な実施態様では、テキストは、少なくとも1つの端末の識別子を含むか、または、テキス
トは、少なくとも1つの端末の識別子およびユーザ名を含む。

【 0 0 1 5 】

第2の態様または第2の態様の第1の可能な実施態様または第2の態様の第2の可能な実施
態様に関連して、第2の態様の第3の可能な実施態様では、本装置は、

割り当てられたセッションが新しいセッションである場合、割り当てられたセッション
と少なくとも2つの端末の識別子との対応関係を確認し、対応関係をデータベースに追加
するように構成された追加ユニットであって、

割り当てられたセッションがデータベース内のセッションである場合、追加された識別
子が割り当てられたセッションに対応するように、少なくとも2つの端末の識別子のうち
にあり、かつデータベースに含まれない端末の識別子をデータベースに追加するようにさ
らに構成された追加ユニット

30

をさらに含む。

【 0 0 1 6 】

第2の態様または第2の態様の第1から第3の可能な実施態様のいずれかに関連して、第2
の態様の第4の可能な実施態様では、本装置は、

第5の端末によって送信された終了メッセージを受信するように構成された第3の受信ユ
ニットと、

データベースから第5の端末の識別子を削除するように構成された削除ユニットであっ
て、

40

セッションに対応する端末の識別子がデータベースにおいて空である場合、データベー
スからセッションを削除するようにさらに構成された削除ユニットと

をさらに含む。

【 0 0 1 7 】

第3の態様によれば、多者通話において音声をテキストに変換するための装置が提供さ
れ、本装置は、

少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するように構
成された第1の受信器であって、音声 - テキスト変換要求が、第1の識別子および第2の識
別子を含む、第1の受信器と、

50

少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てるように構成されたプロセッサと、

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するように構成された第2の受信器であって、多者通話が、1回のセッションに対応する、第2の受信器とを含み、

プロセッサが、音声ストリームをテキストに変換するようにさらに構成され、多者通話において端末にテキストを送信するように構成された送信器をさらに含む。

10

【0018】

第3の態様に関連して、第3の態様の第1の可能な実施態様では、プロセッサは、

データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子に対応する、データベース内のセッションを第1の端末に割り当てることであって、第1の端末が、少なくとも2つの端末のうちのいずれかの端末であり、データベースが、少なくとも1つのセッションおよび該少なくとも1つのセッションに対応する端末識別子を含む、ことと、

データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、かつ第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じである場合、第2の端末および第3の端末に同じ新しいセッションを割り当てることと

20

を行うようにさらに構成される。

【0019】

第3の態様または第3の態様の第1の可能な実施態様に関連して、第3の態様の第2の可能な実施態様では、テキストは、少なくとも1つの端末の識別子を含むか、または、テキストは、少なくとも1つの端末の識別子およびユーザ名を含む。

【0020】

第3の態様または第3の態様の第1の可能な実施態様または第3の態様の第2の可能な実施態様に関連して、第3の態様の第3の可能な実施態様では、

プロセッサは、割り当てられたセッションが新しいセッションである場合、割り当てられたセッションと少なくとも2つの端末の識別子との対応関係を確立し、対応関係をデータベースに追加するようにさらに構成され、

30

プロセッサは、割り当てられたセッションがデータベース内のセッションである場合、追加された識別子が割り当てられたセッションに対応するように、少なくとも2つの端末の識別子のうちにあり、かつデータベースに含まれない端末の識別子をデータベースに追加するようにさらに構成される。

【0021】

第3の態様または第3の態様の第1から第3の可能な実施態様のいずれかに関連して、第3の態様の第4の可能な実施態様では、本装置は、

第5の端末によって送信された終了メッセージを受信するように構成された第3の受信器をさらに含み、

プロセッサは、データベースから第5の端末の識別子を削除するようにさらに構成され、

40

プロセッサは、セッションに対応する端末の識別子がデータベースにおいて空である場合、データベースからセッションを削除するようにさらに構成される。

【0022】

本発明の本実施形態で提供される、多者通話において音声を変換するための

50

方法および装置によれば、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声 - テキスト変換要求が受信され、少なくとも2つの端末によって送信された音声 - テキスト変換要求において同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションが割り当てられ、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8kHzよりも大きい音声ストリームが、パケット交換ドメインを用いて受信され、音声ストリームがテキストに変換され、多者通話において端末にテキストが送信される。受信される音声ストリームは、パケット交換ドメインを用いて受信される、サンプリングレートが8kHzよりも大きい音声ストリームであるため、パケット交換ドメインは、高いサンプリングレートの音声ストリームの送信をサポートすることができ、受信した音声ストリームがテキストに変換された後に取得されるテキストの精度は高くなる。さらに、この処理は、手動操作を必要とせず、比較的迅速であり、人間の労力と時間を節約する。結論として、本発明で提供される方法を用いることによって、多者通話において音声から変換されたテキストを正確に取得することは、時間を節約する容易な方法で実施することができ、これにより、多者通話において音声から変換されたテキストを取得することが時間と労力を要するものであるまたは不正確であるという従来技術の問題が解決されることが学ばれ得る。

10

【0023】

本発明の実施形態または従来技術の技術的解決策をより明確に説明するために、以下では、実施形態または従来技術を説明するために必要な添付図面について簡単に説明する。当然ながら、以下の説明の添付図面は、本発明の一部の実施形態を示しているに過ぎず、当業者は、創造的な労力なしにこれらの添付図面から他の図面をさらに得ることができる。

20

【図面の簡単な説明】

【0024】

【図1】本発明の一実施形態による、多者通話において音声を変換するネットワークアーキテクチャ図である。

【図2】本発明の一実施形態による、多者通話において音声を変換するための方法の概略フローチャートである。

【図3】本発明の一実施形態による、多者通話において音声を変換するための別の方法の概略フローチャートである。

30

【図4A】本発明の一実施形態による、多者通話において音声を変換するためのさらに別の方法の概略フローチャートである。

【図4B】本発明の一実施形態による、多者通話において音声を変換するためのさらに別の方法の概略フローチャートである。

【図5】本発明の一実施形態による、多者通話において音声を変換するための装置の概略構成図である。

【図6】本発明の一実施形態による、多者通話において音声を変換するための別の装置の概略構成図である。

【図7】本発明の一実施形態による、多者通話において音声を変換するためのさらに別の装置の概略構成図である。

40

【図8】本発明の一実施形態による、多者通話において音声を変換するためのさらに別の装置の概略構成図である。

【発明を実施するための形態】

【0025】

以下では、本発明の実施形態の添付図面を参照しながら本発明の実施形態の技術的解決策を明確かつ完全に説明する。当然ながら、説明されている実施形態は、本発明の実施形態の一部に過ぎず、本発明の実施形態の全部ではない。創造的な労力なしに本発明の実施形態に基づいて当業者によって得られる他のすべての実施形態は、本発明の保護範囲内に入るものとする。

50

【0026】

図1は、多者通話において音声テキストに変換するネットワークアーキテクチャ図を示す。端末1～端末Nは、多者通話を確認し、端末1～端末Nはすべて、音声 - テキスト変換機能を有効にする。通話処理において、端末1～端末Nは、パケット交換ドメインを用いて音声ストリームをサーバに送信し、サーバは、音声ストリームをテキストに変換し、サーバは、端末1～端末Nにテキストを送信する。

【0027】

実施形態1

本発明の実施形態1は、多者通話において音声テキストに変換するための方法を提供する。図2に示すように、本方法は、以下のステップを含んでもよい。

【0028】

S201：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信する。

【0029】

音声 - テキスト変換要求は、第1の識別子および第2の識別子を含む。

【0030】

任意選択で、サーバは、オペレータによって用意されたサーバであってもよいし、第三者によって用意されたサーバであってもよい。

【0031】

音声 - テキスト変換要求は、端末に配置されたマイクロホン (Microphone、略してMIC) が有効にされているとき、多者通話中に端末によってサーバに自動的に送信されてもよい。

【0032】

任意選択で、端末に配置されたMICは、以下の2つの方法で有効にされてもよい。

【0033】

方法1：発呼したときまたは呼に応答したときに、ユーザは、MICを手動で有効にする。

【0034】

方法2：ユーザは、ユーザが発呼したときまたは呼に応答したときにデフォルトでMICが有効になるように端末を設定する。

【0035】

実際の適用の過程で、端末のMICを有効にする方法は、実際の要求に応じて設定されてもよく、本発明では限定されないことに留意されたい。

【0036】

音声 - テキスト変換要求は、シグナリングの形式であってもよいし、メッセージの形式であってもよいことに留意されたい。本発明では、音声 - テキスト変換要求の形式は限定されない。音声 - テキスト変換要求が、シグナリングの形式である場合、それは、既存のシグナリングであってもよいし、新たに設定されたシグナリングであってもよい。本発明では、音声 - テキスト変換要求のシグナリングの種類も限定されない。

【0037】

さらに、音声 - テキスト変換要求は、端末のユーザ名および電子メールアドレスなどの他のコンテンツをさらに含んでもよいことにさらに留意されたい。コンテンツは、ユーザによって端末に予め設定されてもよい。実際の適用の過程で、異なるコンテンツが、実際の要求に応じて音声 - テキスト変換要求に追加されてもよい。本発明では、音声 - テキスト変換要求に含まれる特定のコンテンツは限定されない。

【0038】

任意選択で、第1の識別子および第2の識別子は、端末識別子であってもよい。

【0039】

端末識別子は、端末を一意的に特定することができる。

【0040】

好ましくは、端末識別子は、端末の電話番号である。

10

20

30

40

50

【 0 0 4 1 】

多者通話の処理において、多者通話に参加する端末は、発呼端末および被呼端末を含む。発呼端末は、ユーザが能動的に発呼するために使用する端末であり、被呼端末は、ユーザが受動的に呼に応答するために使用する端末である。

【 0 0 4 2 】

具体的には、第1の識別子は、発呼端末の識別子であってもよく、これに対応して、第2の識別子は、被呼端末の識別子であってもよく、あるいは、第1の識別子は、被呼端末の識別子であってもよく、これに対応して、第2の識別子は、発呼端末の識別子であってもよい。

【 0 0 4 3 】

例えば、電話番号が123の端末1と、電話番号が456の端末2が通話中だと仮定して（端末1が発呼端末であり、端末2が被呼端末であると仮定して）、端末1が、音声 - テキスト変換機能を有効にしたとき、サーバは、端末1の音声 - テキスト変換要求 < 123, 456 > を受信し、端末2が、音声 - テキスト変換機能を有効にしたとき、サーバは、端末2の音声 - テキスト変換要求 < 123, 456 > を受信する。

【 0 0 4 4 】

S202：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションを割り当てる。

【 0 0 4 5 】

セッションは、複数の通話者（2人の通話者を含む）が電話通信を行うプロセスである。多者通話が完全に確立され、多者通話に参加する各通話者が通話をテキストに変換する機能を有効にしたとき、通話に参加するすべての通話者によって使用される端末に、セッションが割り当てられる。

【 0 0 4 6 】

例えば、多者通話が、10人の参加者を含み、10人の参加者すべてが、通話をテキストに変換する機能を有効にする場合、セッションは、10人の参加者が使用する端末に割り当てられる。

【 0 0 4 7 】

サーバが、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションを割り当てることは、以下の3つのケースを含み得る。

【 0 0 4 8 】

ケース1：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子および同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションを割り当てる。

【 0 0 4 9 】

例えば、サーバは、端末Aの音声 - テキスト変換要求1 < 123, 456 > を受信し、端末Bの音声 - テキスト変換要求2 < 123, 456 > を受信する。端末Aの音声 - テキスト変換要求に含まれる第1の識別子123は、端末Bの音声 - テキスト変換要求に含まれる第1の識別子123と同じであり、端末Aの音声 - テキスト変換要求に含まれる第2の識別子456は、端末Bの音声 - テキスト変換要求に含まれる第2の識別子456と同じであるため、サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子および同じ第2の識別子を有する端末に同じセッションが割り当てられるように、端末Aおよび端末Bに同じセッションを割り当てる。

【 0 0 5 0 】

ケース2：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てる。

10

20

30

40

50

【 0 0 5 1 】

例えば、サーバは、端末1によってサーバに送信された音声 - テキスト変換要求1 < 100 , 001 > を受信し、端末2の音声 - テキスト変換要求2 < 100 , 002 > を受信する。音声 - テキスト変換要求1および音声 - テキスト変換要求2は、同じ第1の識別子 (100) を有する。したがって、同じセッションが、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子を有する端末が同じセッションを有するように、端末1および端末2に割り当てられる。

【 0 0 5 2 】

ケース3：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てる。

10

【 0 0 5 3 】

例えば、サーバは、端末1によってサーバに送信された音声 - テキスト変換要求1 < 001 , 100 > を受信し、端末2の音声 - テキスト変換要求2 < 002 , 100 > を受信する。音声 - テキスト変換要求1および音声 - テキスト変換要求2は、同じ第2の識別子 (100) を有する。したがって、同じセッションが、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第2の識別子を有する端末が同じセッションを有するように、端末1および端末2に割り当てられる。

【 0 0 5 4 】

S203：サーバは、1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信する。

20

【 0 0 5 5 】

1回の多者通話は、1回のセッションに対応する。

【 0 0 5 6 】

パケット交換ドメインは、データサービスを送信するためのリンクを指す。パケット交換ドメインを用いて送信される音声のサンプリングレートは、8KHzより大きくてもよい。

【 0 0 5 7 】

サンプリングレートは、ヘルツ (Hertz、略してHz) で測定されるサンプリング周波数、すなわち、1秒当たりの、連続信号から取り出されるサンプリング量とも呼ばれ得る。

30

【 0 0 5 8 】

好ましくは、セッションに参加する端末によって送信され、パケット交換ドメインを用いてサーバによって受信されるサンプリングレートは、16キロヘルツ (KiloHertz、略してKHz) である。

【 0 0 5 9 】

サーバが、セッションに参加する端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信する前に、端末とサーバとの間で、データ接続を確立する必要があることに留意されたい。任意選択で、データ接続を確立する方法は、伝送制御プロトコル (Transmission Control Protocol、略してTCP) ソケット接続およびハイパーテキスト転送プロトコル (Hypertext Transfer Protocol、略してHTTP) 接続などを含むが、これらに限定されない。前述のデータ接続方法を用いることによって、端末によってサーバに送信される音声ストリームは、パケット交換ドメインにおいて送信される。

40

【 0 0 6 0 】

S204：サーバは、音声ストリームをテキストに変換する。

【 0 0 6 1 】

任意選択で、サーバにインストールされた音声認識エンジンが、音声ストリームを通話テキストに変換してもよい。

【 0 0 6 2 】

好ましくは、通話テキストへの音声ストリームの変換は、リアルタイムで実行され得る

50

。【0063】

サーバによって音声ストリームから変換されたテキストは、多者通話の各参加者の音声ストリームから変換されたテキストであり、テキストに含まれるコンテンツは、多者通話の各参加者の発話の順番に従ってソートされることに留意されたい。

【0064】

例えば、多者通話は、3人の参加者を含み、3人の参加者は、それぞれ参加者1、参加者2、および参加者3である。多者通話において、参加者1が、「議論にはどこが比較的適していると思いますか」と発言し、次に参加者2が、「第1会議室が比較的適していると思います」と発言し、次に参加者3が、「本社が比較的適していると思います」と発言すると仮定しよう。このとき、サーバによって多者通話の音声ストリームから変換されたテキストは、以下の通りとなる。

「議論にはどこが比較的適していると思いますか。」

第1会議室が比較的適していると思います。

本社が比較的適していると思います。」

【0065】

S205：サーバは、多者通話において端末にテキストを送信する。

【0066】

例えば、多者通話の参加者は、端末1、端末2、および端末3を含むと仮定しよう。サーバは、端末1、端末2、および端末3の音声ストリームを受信し、受信した3つの端末の音声ストリームをテキストに変換し、端末1、端末2、および端末3に送信する。

【0067】

任意選択で、サーバは、各セッションに一意のセッション識別子を割り当てることによって異なるセッションを区別してもよく、サーバが少なくとも2つの端末にセッションを割り当てるとき、セッションに割り当てられたセッション識別子をセッションの端末に送信してもよい。

【0068】

これに対応して、サンプリングレートが8KHzよりも大きく、1回の多者通話において少なくとも1つの端末によって送信され、パケット交換ドメインを用いてサーバによって受信される音声ストリームも、サーバがセッションを区別するためにセッション識別子を含んでもよい。

【0069】

本発明の本実施形態で提供される、多者通話において音声を変換するための方法によれば、サーバは、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声 - テキスト変換要求を受信し、少なくとも2つの端末によって送信された音声 - テキスト変換要求において同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションを割り当て、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信し、音声ストリームをテキストに変換し、多者通話において端末にテキストを送信する。サーバによって受信される音声ストリームは、パケット交換ドメインを用いて受信される、サンプリングレートが8KHzよりも大きい音声ストリームであるため、パケット交換ドメインは、高いサンプリングレートの音声ストリームの送信をサポートすることができ、受信した音声ストリームがテキストに変換された後に取得されるテキストの精度は高くなる。さらに、この処理は、手動操作を必要とせず、比較的迅速であり、人間の労力と時間を節約する。結論として、本発明で提供される方法を用いることによって、多者通話において音声から変換されたテキストを正確に取得することは、時間を節約する容易な方法で実施することができ、これにより、多者通話において音声から変換されたテキストを取得することが時間と労力を要するものであるまたは不正確であるという従来技術の問題が解決されることが学ばれ得る。

10

20

30

40

50

【0070】

実施形態2

本発明の実施形態2は、多者通話において音声をテキストに変換するための方法を提供する。図3に示すように、本方法は、以下のステップを含んでもよい。

【0071】

S301：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信する。

【0072】

音声 - テキスト変換要求は、第1の識別子および第2の識別子を含む。

【0073】

S301はS201と同じであり、本明細書では詳細を説明しないことに留意されたい。

【0074】

S302：サーバは、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てる。

【0075】

具体的には、サーバが少なくとも2つの端末にセッションを割り当てることは、以下の2つのケースを含む。

【0076】

ケース1：データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じである場合、同じ新しいセッションが、第2の端末および第3の端末に割り当てられる。

【0077】

データベースは、少なくとも1つのセッションおよび少なくとも1つのセッションに対応する端末識別子を含む。

【0078】

任意選択で、データベースは、表の形式で記憶されてもよい。表1は、データベースの種類を示している。

【0079】

【表1】

表 1

セッション	端末識別子
セッション 1	端末 1 および 端末 2
セッション 2	端末 3、端末 4、および 端末 5
セッション 3	端末 8、端末 9、端末 10、および 端末 11
...	...

【0080】

表1は、データベースの形式およびコンテンツを単に表の形式で示したものであり、データベースの形式およびコンテンツを特に限定するものではないことに留意されたい。もちろん、データベースは、別の形式で記憶されてもよく、他のコンテンツを含んでもよく、本発明では限定されない。

10

20

30

40

50

【 0 0 8 1 】

例1：電話番号が111の端末1および電話番号が222の端末2が通話を確立し、端末1および端末2の両方が、通話をテキストに変換する機能を有効にする。サーバは、端末1の音声 - テキスト変換要求1 < 111 , 222 > および端末2の音声 - テキスト変換要求2 < 111 , 222 > を受信する。

【 0 0 8 2 】

データベースは、端末1および端末2によって送信された音声 - テキスト変換要求の第1の識別子(111)および第2の識別子(222)を含まないと仮定しよう。音声 - テキスト変換要求1の第1の識別子(111)は、音声 - テキスト変換要求2の第1の識別子(111)と同じであり、音声 - テキスト変換要求1の第2の識別子(222)は、音声 - テキスト変換要求2の第2の識別子(222)と同じであるため、同じ新しいセッション4が、端末1および端末2に割り当てられる。

10

【 0 0 8 3 】

ケース2：データベースは、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含み、

第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子に対応する、データベース内のセッションが、第1の端末に割り当てられる。

【 0 0 8 4 】

第1の端末は、少なくとも2つの端末のうちのいずれかの端末である。

【 0 0 8 5 】

例2：ケース1の例1を踏まえると、端末1と端末2との間の通話中に、端末1が、電話番号が333の端末3に発呼し、これにより、端末3が、端末1と端末2との間の多者通話に参加する。端末1および端末3の両方が通話をテキストに変換する機能を有効にすると仮定しよう。サーバは、端末1の音声 - テキスト変換要求3 < 111 , 333 > および端末3の音声 - テキスト変換要求4 < 111 , 333 > を受信する。

20

【 0 0 8 6 】

データベースは、音声 - テキスト変換要求3および音声 - テキスト変換要求4の第1の識別子(111)を含み、第1の識別子(111)に対応するセッションが、セッション4であるため、第1の識別子(111)に対応する、データベース内のセッション4が、端末1および端末3に割り当てられる。

30

【 0 0 8 7 】

多者通話では、共通端末が複数の参加者に発呼する方法で多者通話が確立されるか、または複数の参加者が共通端末に発呼する方法で多者通話が確立されると仮定されていることに留意されたい。複数の参加者との通話を確立した後、共通端末は、デフォルトで音声 - テキスト変換要求をサーバに送信する。さらに、多者通話の処理において、共通端末は、参加者との通話を最初に確立するとき一度だけ音声 - テキスト変換要求をサーバに送信してもよく、この場合、サーバが共通端末にセッションを割り当てる必要があるのも同様に一度だけである。

【 0 0 8 8 】

例えば、端末1(識別子が111である)が、共通端末(識別子が100である)との多者通話を確立した後、端末1は、音声 - テキスト変換要求1 < 111 , 000 > をサーバに送信し、共通端末は、音声 - テキスト変換要求2 < 111 , 000 > をサーバに送信し、サーバは、端末1および共通端末にセッション1を割り当てる。多者通話において、端末2(識別子が111)は、共通端末(識別子が100である)に発呼することによって多者通話に参加する。次に、端末2は、音声 - テキスト変換要求1 < 222 , 100 > をサーバに送信するが、共通端末は、もはや音声 - テキスト変換要求をサーバに送信しない。この場合、サーバは、単にセッション1を端末2に割り当て、もはや共通端末にセッションを割り当てない。

40

【 0 0 8 9 】

音声 - テキスト変換要求はミーティング・フラグ・ビットをさらに含んでもよいことにさらに留意されたい。ミーティング・フラグ・ビットは、当該通話の参加者が2人または

50

複数の通話者（2人より多い通話者）であることを示すために使用される。第2の端末によって送信された音声 - テキスト変換要求の第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じであり、ミーティング・フラグ・ビットが、当該通話の参加者が2人の通話者であることを示す場合、サーバは、第2の端末および第3の端末にセッションを割り当てるときに、データベースが第2の端末および第3の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含むか否かを判定する必要がなく、第2の端末および第3の端末に新しいセッションを直接割り当ててもよい。

【0090】

10

任意選択で、発呼するときにユーザは、ミーティング・フラグ・ビットによって示されるコンテンツを手動で選択してもよいし、あるいは、ユーザは、ある電話番号を、参加者が複数の通話者（2人より多い通話者）である種類に設定してもよく、ユーザが、この電話番号をダイヤルするとき、参加者が複数の通話者（2人より多い通話者）であることを示すミーティング・フラグ・ビットのコンテンツがデフォルトで選択される。

【0091】

例えば、電話番号100が、共通のミーティング電話番号であり、ユーザが、この電話番号を、参加者が複数の通話者（2人より多い通話者）である種類に設定すると仮定しよう。電話番号をダイヤルするとき、ユーザは、ミーティング・フラグ・ビットによって示されるコンテンツを手動で選択する必要はなく、音声 - テキスト変換要求に含まれるミーティング・フラグ・ビットは、参加者が複数の通話者（2人より多い通話者）であることを自動的に示す。

20

【0092】

S303：サーバは、割り当てられたセッションに従ってデータベースを更新する。

【0093】

具体的には、サーバが割り当てられたセッションに従ってデータベースを更新する処理は、以下の2つのケースを含み得る。

【0094】

ケース1：割り当てられたセッションは、新しいセッションであり、サーバは、割り当てられたセッションと少なくとも2つの端末の識別子との対応関係を確立し、この対応関係をデータベースに追加する。

30

【0095】

ケース2：割り当てられたセッションが、データベース内のセッションであり、少なくとも2つの端末の識別子のうちにあり、かつデータベースに含まれない端末識別子が、追加された識別子が割り当てられたセッションに対応するようにデータベースに追加される。

【0096】

例えば、サーバによって端末12および端末13に割り当てられたセッションが、新しいセッション4であると仮定しよう。このとき、サーバは、セッション4と、端末12および端末13の識別子との対応関係を確立し、この対応関係をデータベースに追加する。元のデータベースの形式が、表1に示されているとした場合、この対応関係がデータベースに追加された後では、データベースは、表2に示すものとなる。

40

【0097】

【表 2】

表 2

セッション	端末識別子
セッション 1	端末 1 および端末 2
セッション 2	端末 3、端末 4、および端末 5
セッション 3	端末 8、端末 9、端末 10、および端末 11
セッション 4	端末 12 および端末 13
...	...

10

【 0 0 9 8 】

さらに、サーバによって端末12および端末14に割り当てられたセッションが、データベース内のセッション4であると仮定しよう。データベースは、端末12の識別子を含むが、端末14の識別子を含まないため、サーバは、追加された識別子が割り当てられたセッションに対応するように、データベースに含まれていない端末識別子（端末14の）をデータベースに追加する。端末14の識別子が追加されたデータベースを表3に示す。

【 0 0 9 9 】

20

【表 3】

表 3

セッション	端末識別子
セッション 1	端末 1 および端末 2
セッション 2	端末 3、端末 4、および端末 5
セッション 3	端末 8、端末 9、端末 10、および端末 11
セッション 4	端末 12、端末 13、および端末 14
...	...

30

【 0 1 0 0 】

S304：サーバは、1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信する。

【 0 1 0 1 】

この多者通話は、1回のセッションに対応する。

【 0 1 0 2 】

S304はS203と同じであり、本明細書では詳細を説明しないことに留意されたい。

40

【 0 1 0 3 】

多者通話における少なくとも1つの端末が、パケット交換ドメインを用いて、サンプリングレートが8KHzよりも大きい音声ストリームをサーバに送信する前に、さらに、端末とサーバとの間でデータ接続を確立する必要があることにさらに留意されたい。

【 0 1 0 4 】

任意選択で、端末とサーバとの間のデータ接続の確立は、通常、端末がデータ接続要求を開始する方法で実施される。端末によって開始されたデータ接続要求は、S301で音声 - テキスト変換要求と組み合わせられ、1つの要求にされてもよい。さらに、データ接続は、端末とサーバとの間で確立される。データ接続を確立する方法は、伝送制御プロトコルTC

50

Pソケット接続およびHTTP接続などを含むが、これらに限定されない。前述のデータ接続を用いることによって、サーバによって受信される音声ストリームは、パケット交換ドメインにおいて送信され得る。

【0105】

もちろん、データ接続要求および音声-テキスト変換要求は、2つの独立した要求であってもよく、本発明では限定されない。

【0106】

S303およびS304は、順番に実行されてもよいし、同時に実行されてもよく、このことは、本発明では限定されないことにさらに留意されたい。

【0107】

S305：サーバは、音声ストリームをテキストに変換する。

【0108】

任意選択で、テキストは、少なくとも1つの端末の識別子を含んでもよいし、テキストは、少なくとも1つの端末の識別子およびユーザ名を含んでもよい。

【0109】

例えば、ユーザAは、電話番号が111（端末の識別子が111）の端末の端末ユーザ名を「Xiao Ming」に設定し、ユーザBは、電話番号が222（端末の識別子が222）の端末の端末ユーザ名を「Xiao Hong」に設定する。S301において、端末によってサーバに送信される音声-テキスト変換要求は、端末ユーザ名を含む。ユーザAとユーザBとの間の通話中、ユーザAは、「どこで契約書に署名しましょうか」と発言し、ユーザBは、「あなたの会社で契約書に署名しましょう」と発言する。このとき、この通話中に、サーバによって音声ストリームから変換されたテキストは、以下の通りとなる。

「Xiao Ming (111)：どこで契約書に署名しましょうか。」

Xiao Hong (222)：あなたの会社で契約書に署名しましょう。」

【0110】

S306：サーバは、多者通話において端末にテキストを送信する。

【0111】

任意選択で、サーバは、多者通話においてリアルタイムで端末にテキストを送信してもよいし、予め設定された期間に従ってテキストを送信してもよいし、多者通話において端末によって送信された終了メッセージを受信した後にテキストを送信してもよい。

【0112】

具体的には、サーバが多者通話において端末にテキストを送信する契機は、以下の4つのケースを含み得るが、これらに限定されない。

【0113】

ケース1：サーバは、予め設定された期間に従って、多者通話においてこの期間のテキストを端末に送信する。

【0114】

実際の適用の過程で、予め設定された期間は、実際の要求に応じて設定されてもよく、本発明では限定されないことに留意されたい。

【0115】

例えば、多者通話における端末が、端末1および端末2であり、予め設定された期間が、1分間であると仮定しよう。このとき、端末1および端末2が通話を開始してから、1分毎に、サーバは、当該の1分間の2人の通話者間の通話のテキストを端末1および端末2に送信する。

【0116】

ケース2：多者通話においてすべての端末によって送信された終了メッセージを受信した後、サーバは、多者通話においてすべての端末にテキストを送信する。

【0117】

終了メッセージは、通話の終了を示すメッセージである。

【0118】

10

20

30

40

50

任意選択で、終了メッセージは、オンフック要求であってもよい。

【0119】

例えば、多者通話における端末が、端末3、端末4、および端末5であると仮定しよう。3つの端末が同時に通話を開始し、最初に端末3がサーバに終了メッセージを送信し、次に端末4および端末5がサーバに終了メッセージを送信すると仮定しよう。3つの端末すべてが、サーバに終了メッセージを送信した後（端末4および端末5が、サーバに終了メッセージを送信した後）、サーバは、この多者通話のテキストを3つの端末の各端末に送信する。

【0120】

ケース3：多者通話においてすべての端末によって送信された終了メッセージを受信した後、サーバは、多者通話において、各端末が参加したプロセスのテキストを各端末に送信する。

10

【0121】

例えば、多者通話における端末が、端末6、端末7、および端末8であると仮定しよう。3つの端末が同時に通話を開始すると仮定しよう。次に、3分間の通話の後に、端末6がサーバに終了メッセージを送信し、5分間の通話の後に、端末7および端末8がサーバに終了メッセージを送信する。

【0122】

端末6、端末7、および端末8によって送信された終了メッセージを受信した後、サーバは、端末6が参加した、端末6、端末7、および端末8間の3分間の通話のテキストを端末6に送信し、端末7が参加した、端末6、端末7、および端末8間の5分間の通話のテキストを端末7に送信し、端末8が参加した、端末6、端末7、および端末8間の5分間の通話のテキストを端末8に送信する。

20

【0123】

ケース4：多者通話において第1の端末によって送信された終了メッセージを受信した後、サーバは、この端末が参加したプロセスのテキストを第1の端末にのみ送信する。

【0124】

例えば、多者通話における端末が、端末9、端末10、および端末11であると仮定しよう。3つの端末が同時に通話を開始すると仮定しよう。次に、3分間の通話の後に、端末9がサーバに終了メッセージを送信し、5分間の通話の後に、端末10および端末11がサーバに終了メッセージを送信する。

30

【0125】

端末9によって送信された終了メッセージを受信した後、サーバは、端末9が参加した、端末9、端末10、および端末11間の3分間の通話のテキストを端末9に送信する。

【0126】

端末10によって送信された終了メッセージを受信した後、サーバは、端末10が参加した、端末9、端末10、および端末11間の5分間の通話のテキストを端末10に送信する。

【0127】

端末11によって送信された終了メッセージを受信した後、サーバは、端末11が参加した、端末9、端末10、および端末11間の5分間の通話のテキストを端末11に送信する。

40

【0128】

さらに、ケース4では、端末が参加したプロセスの通話のテキストが第1の端末に送信された後に、この方法は、

第1の端末に対応するセッションに対応する、データベース内のすべての端末に、第1の端末が参加したプロセスの、各端末の通話のテキストを送信することをさらに含んでもよい。

【0129】

例えば、多者通話における端末が、端末12および端末13であると仮定しよう。サーバが、端末12によって送信された終了メッセージを受信すると、サーバは、端末12が参加したプロセスの、端末12と端末13との間の通話のテキストを端末12および端末13に送信する。

50

【0130】

前述の4つのケースは、サーバが多者通話において端末にテキストを送信する契機を単なる例の形で説明しているが、サーバが多者通話において端末にテキストを送信する契機を限定しないことに留意されたい。実際の適用の過程で、サーバが多者通話において端末にテキストを送信する契機は、実際の要求に応じて設定されてもよく、本発明では限定されないことに留意されたい。

【0131】

S307：サーバは、第5の端末によって送信された終了メッセージを受信する。

【0132】

S308：サーバは、データベースから第5の端末の識別子を削除する。

10

【0133】

例えば、データベースにおいて、セッション1に対応する端末が、端末1、端末2、および端末3であると仮定しよう。サーバは、端末1によって送信された終了メッセージ（通話を切る）を受信し、次に、サーバは、端末1の識別子をデータベースから削除する。

【0134】

S309：セッションに対応する端末識別子が、データベースにおいて空である場合、サーバは、データベースからセッションを削除する。

【0135】

例えば、セッション1に対応する端末識別子が、データベースにおいて空であると仮定しよう。このとき、サーバは、データベースからセッション1を削除する。

20

【0136】

本発明の本実施形態で提供される、多者通話において音声を変換するための方法によれば、サーバは、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声-テキスト変換要求を受信し、少なくとも2つの端末によって送信された音声-テキスト変換要求において同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションを割り当て、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信し、音声ストリームをテキストに変換し、多者通話において端末にテキストを送信する。サーバによって受信される音声ストリームは、パケット交換ドメインを用いて受信される、サンプリングレートが8KHzよりも大きい音声ストリームであるため、パケット交換ドメインは、高いサンプリングレートの音声ストリームの送信をサポートすることができ、受信した音声ストリームがテキストに変換された後に取得されるテキストの精度は高くなる。さらに、この処理は、手動操作を必要とせず、比較的迅速であり、人間の労力と時間を節約する。結論として、本発明で提供される方法を用いることによって、多者通話において音声から変換されたテキストを正確に取得することは、時間を節約する容易な方法で実施することができ、これにより、多者通話において音声から変換されたテキストを取得することが時間と労力を要するものであるまたは不正確であるという従来技術の問題が解決されることが学ばれ得る。

30

【0137】

実施形態3

本発明の実施形態3は、多者通話において音声を変換するための方法を提供する。

40

【0138】

Zhang Sanが端末Aを使用して、端末Bを使用するLi Siに発呼し、端末Aの電話番号が123であり、端末Bの電話番号が456であり、端末AでZhang Sanによって設定された、端末Aのユーザ名がZhang Sanであり、端末BでLi Siによって設定された、端末Bのユーザ名がLi Siであると仮定しよう。

【0139】

図3の方法を説明するために、サーバ、端末A、および端末B間の相互作用プロセスを例

50

として使用する。図4Aおよび図4Bに示すように、この方法は、以下のステップを含んでもよい。

【 0 1 4 0 】

S401：端末Aは、音声 - テキスト変換要求1をサーバに送信する。

【 0 1 4 1 】

例えば、端末Aが端末Bに発呼した後、Zhang Sanは、MICを有効にし、この通話が参加者が2人より多い通話者である通話であることを示すミーティング・フラグ・ビットを選択する。次に、端末Aは、Zhang San（電話番号が123である端末）が電話番号が456である端末電話に発呼し、通話の参加者が2人より多い通話者であることを示す音声 - テキスト変換要求1 < 123, Zhang San, 456, TRUE > をサーバに送信する。

10

【 0 1 4 2 】

S402：端末Bは、音声 - テキスト変換要求2をサーバに送信する。

【 0 1 4 3 】

例えば、Li Siは、Zhang Sanによる発呼に応答し、MICを有効にする。次に、端末Bは、電話番号が123である端末が電話番号が456である端末に発呼したことを示す音声 - テキスト変換要求2 < 123, 456, Li Si > をサーバに送信する。

【 0 1 4 4 】

S403：サーバは、音声 - テキスト変換要求1および音声 - テキスト変換要求2に含まれる第1の識別子が同じであり、音声 - テキスト変換要求1および音声 - テキスト変換要求2に含まれる第2の識別子が同じであることを判定し、音声 - テキスト変換要求1に含まれるミーティング・フラグ・ビットが示す参加者が2人より多い通話者であるか否かをさらに判定する。

20

【 0 1 4 5 】

音声 - テキスト変換要求1に含まれるミーティング・フラグ・ビットが示す参加者が、2人より多い通話者ではない場合、S404が実行される。

【 0 1 4 6 】

音声 - テキスト変換要求1に含まれるミーティング・フラグ・ビットが示す参加者が、2人より多い通話者である場合、S405が実行される。

【 0 1 4 7 】

S404：サーバは、端末Aおよび端末Bに新しいセッションを割り当てる。

30

【 0 1 4 8 】

例えば、サーバは、端末Aおよび端末Bに新しいセッション1を割り当てる。

【 0 1 4 9 】

S405：サーバは、データベースが音声 - テキスト変換要求1および音声 - テキスト変換要求2の第1の識別子または第2の識別子を含むか否かを判定する。

【 0 1 5 0 】

データベースが、音声 - テキスト変換要求1および音声 - テキスト変換要求2の第1の識別子または第2の識別子を含まない場合、S406が実行される。

【 0 1 5 1 】

データベースが、音声 - テキスト変換要求1および音声 - テキスト変換要求2の第1の識別子または第2の識別子を含む場合、S408が実行される。

40

【 0 1 5 2 】

S406：サーバは、端末Aおよび端末Bに新しいセッションを割り当てる。

【 0 1 5 3 】

S404およびS406の後に、S407は実行される必要があることに留意されたい。

【 0 1 5 4 】

S407：サーバは、割り当てられたセッションと、端末Aの識別子および端末Bの識別子との対応関係を確立し、この対応関係をデータベースに追加する。

【 0 1 5 5 】

例えば、S404またはS406において、端末Aおよび端末Bに割り当てられたセッションが、

50

セッション1であると仮定しよう。このとき、セッション1と、端末Aの識別子(123)および端末Bの識別子(456)との対応関係が確立され、この対応関係が、データベースに追加される。

【0156】

S408：サーバは、端末Aおよび端末Bによって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子に対応する、データベース内のセッションを端末Aおよび端末Bに割り当てる。

【0157】

例えば、サーバは、データベースが、音声 - テキスト変換要求1および音声 - テキスト変換要求2の第2の識別子456(端末Bの識別子)を含むと判定し、データベースにおいて、第2の識別子456に対応するセッションは、セッション2である。次に、セッション2が、端末Aおよび端末Bに割り当てられる。

10

【0158】

S409：サーバは、追加された識別子が割り当てられたセッションに対応するように、端末Aの識別子および端末Bの識別子のうちにあるがデータベースに含まれていない端末識別子をデータベースに追加する。

【0159】

例えば、サーバは、データベースが、音声 - テキスト変換要求1および音声 - テキスト変換要求2の第2の識別子456(端末Bの識別子)を含むが、第1の識別子123(端末Aの識別子)を含まないと判定し、データベースにおいて、第2の識別子456に対応するセッションは、セッション2である。次に、サーバは、追加された識別子がセッション2に対応するように、端末Aの識別子をデータベースに追加する。

20

【0160】

S410：端末Aは、パケット交換ドメインを用いて、サンプリングレートが8KHzよりも大きい音声ストリームをサーバに送信する。

【0161】

例えば、端末Aは、パケット交換ドメインを用いて、サンプリングレートが8KHzよりも大きい、「どこで会議を行いましょうか」という音声ストリームをサーバに送信する。

【0162】

S411：端末Bは、パケット交換ドメインを用いて、サンプリングレートが8KHzよりも大きい音声ストリームをサーバに送信する。

30

【0163】

例えば、端末Bは、パケット交換ドメインを用いて、サンプリングレートが8KHzよりも大きい、「本社の第1会議室で行いましょう」という音声ストリームをサーバに送信する。

【0164】

S412：サーバは、端末Aおよび端末Bによって送信された音声ストリームをテキストに変換する。

【0165】

例えば、S410およびS411の例を踏まえると、サーバは、端末Aによって送信された音声ストリームを以下のようなテキストに変換する。

40

Zhang San(123)：どこで会議を行いましょうか。

Li Si(456)：本社の第1会議室で行いましょう。

【0166】

S413：端末Aは、サーバに終了メッセージを送信する。

【0167】

例えば、通話を切った後、端末Aは、終了メッセージ「通信終了(OVER)」をサーバに送信する。

【0168】

S414：サーバは、端末Aにテキストを送信する。

50

【 0 1 6 9 】

例えば、S412の例を踏まえると、サーバは、以下のようなテキストを端末Aに送信する

Zhang San (123) : どこで会議を行いましょうか。

Li Si (456) : 本社の第1会議室で行いましょう。

【 0 1 7 0 】

S415 : サーバは、データベースから端末Aの識別子を削除する。

【 0 1 7 1 】

S416 : 端末Bは、サーバに終了メッセージを送信する。

【 0 1 7 2 】

例えば、通話を切った後、端末Bは、終了メッセージ「通信終了」をサーバに送信する

【 0 1 7 3 】

S417 : サーバは、端末Bにテキストを送信する。

【 0 1 7 4 】

例えば、S412の例を踏まえると、サーバは、以下のようなテキストを端末Bに送信する

Zhang San (123) : どこで会議を行いましょうか。

Li Si (456) : 本社の第1会議室で行いましょう。

【 0 1 7 5 】

S418 : サーバは、データベースから端末Bの識別子を削除する。

【 0 1 7 6 】

S419 : サーバは、セッションに対応する端末の識別子がデータベースにおいて空であると判定し、データベースからセッションを削除する。

【 0 1 7 7 】

例えば、サーバは、セッション3に対応する端末の識別子が空であると判定し、次に、サーバは、データベースからセッション3を削除する。

【 0 1 7 8 】

本発明の本実施形態で提供される、多者通話において音声を変換するための方法によれば、サーバは、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声 - テキスト変換要求を受信し、少なくとも2つの端末によって送信された音声 - テキスト変換要求において同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションを割り当て、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信し、音声ストリームをテキストに変換し、多者通話において端末にテキストを送信する。サーバによって受信される音声ストリームは、パケット交換ドメインを用いて受信される、サンプリングレートが8KHzよりも大きい音声ストリームであるため、パケット交換ドメインは、高いサンプリングレートの音声ストリームの送信をサポートすることができ、受信した音声ストリームがテキストに変換された後に取得されるテキストの精度は高くなる。さらに、この処理は、手動操作を必要とせず、比較的迅速であり、人間の労力と時間を節約する。結論として、本発明で提供される方法を用いることによって、多者通話において音声から変換されたテキストを正確に取得することは、時間を節約する容易な方法で実施することができ、これにより、多者通話において音声から変換されたテキストを取得することが時間と労力を要するものであるまたは不正確であるという従来技術の問題が解決されることが学ばれ得る。

【 0 1 7 9 】

実施形態4

本発明の実施形態4は、多者通話において音声を変換するための装置50を提供する。装置50は、前述の実施形態ではサーバの一部または全部である。図5に示すよう

10

20

30

40

50

に、多者通話において音声を変換するための装置50は、

少なくとも2つの端末によって送信された音声 - テキスト変換要求を受信するように構成された第1の受信ユニット501であって、音声 - テキスト変換要求が、第1の識別子および第2の識別子を含む、第1の受信ユニット501と、

少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てるように構成された割り当てユニット502と、

1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するように構成された第2の受信ユニット503であって、多者通話が、1回のセッションに対応する、第2の受信ユニット503と、

音声ストリームを変換するように構成された変換ユニット504と、

多者通話において端末にテキストを送信するように構成された送信ユニット505とを含んでもよい。

【0180】

さらに、割り当てユニット502は、

第1の端末が、少なくとも2つの端末のうちのいずれかの端末であり、データベースが、少なくとも1つのセッションおよび少なくとも1つのセッションに対応する端末識別子を含むとき、データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子に対応する、データベース内のセッションを第1の端末に割り当て、

データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じである場合、第2の端末および第3の端末に同じ新しいセッションを割り当てる

ように特に構成されてもよい。

【0181】

任意選択で、テキストは、少なくとも1つの端末の識別子を含んでもよいし、テキストは、少なくとも1つの端末の識別子およびユーザ名を含んでもよい。

【0182】

さらに、図6に示すように、多者通話において音声を変換するための装置50は、

割り当てられたセッションが新しいセッションである場合、割り当てられたセッションと少なくとも2つの端末の識別子との対応関係を確認し、この対応関係をデータベースに追加するように構成された追加ユニット506であって、

割り当てられたセッションがデータベース内のセッションである場合、追加された識別子が割り当てられたセッションに対応するように、少なくとも2つの端末の識別子のうち

にあり、かつデータベースに含まれない端末識別子をデータベースに追加するようにさらに構成されてもよい追加ユニット506

【0183】

さらに、図6に示すように、多者通話において音声を変換するための装置50は、

第5の端末によって送信された終了メッセージを受信するように構成された第3の受信ユニット507と、

データベースから第5の端末の識別子を削除するように構成された削除ユニット508であって、

10

20

30

40

50

セッションに対応する端末の識別子がデータベースにおいて空である場合、データベースからセッションを削除するようにさらに構成されてもよい削除ユニット508とをさらに含んでもよい。

【0184】

本発明の本実施形態で提供される、多者通話において音声を変換するための装置50によれば、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声 - テキスト変換要求が受信され、少なくとも2つの端末によって送信された音声 - テキスト変換要求において同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションが割り当てられ、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームが、パケット交換ドメインを用いて受信され、音声ストリームがテキストに変換され、多者通話において端末にテキストが送信される。受信される音声ストリームは、パケット交換ドメインを用いて受信される、サンプリングレートが8KHzよりも大きい音声ストリームであるため、パケット交換ドメインは、高いサンプリングレートの音声ストリームの送信をサポートすることができ、受信した音声ストリームがテキストに変換された後に取得されるテキストの精度は高くなる。さらに、この処理は、手動操作を必要とせず、比較的迅速であり、人間の労力と時間を節約する。結論として、本発明で提供される方法を用いることによって、多者通話において音声から変換されたテキストを正確に取得することは、時間を節約する容易な方法で実施することができ、これにより、多者通話において音声から変換されたテキストを取得することが時間と労力を要するものであるまたは不正確であるという従来技術の問題が解決されることが学ばれ得る。

【0185】

実施形態5

本発明の実施形態5は、多者通話において音声を変換するための装置50を提供する。装置50は、前述の実施形態ではサーバの一部または全部である。図7に示すように、多者通話において音声を変換するための装置50は、

少なくとも1つのプロセッサ701と、装置間の接続および相互通信を実施するように構成された少なくとも1つの通信バス702と、第1の受信器703と、第2の受信器704と、送信器705とを含んでもよい。

【0186】

通信バス702は、業界標準アーキテクチャ (Industry Standard Architecture、略してISA) バス、周辺コンポーネント相互接続 (Peripheral Component Interconnect、略してPCI) バス、または拡張業界標準アーキテクチャ (Extended Industry Standard Architecture、略してEISA) バスなどであってもよい。バスは、アドレスバス、データバス、および制御バスなどに分類され得る。表示の都合上、バスは、図7では、単に1本の太い線を用いて示されているが、このことは、1つのバスまたは1つの種類のバスしか存在しないことを示していない。

【0187】

プロセッサ701は、中央処理装置 (Central Processing Unit、略してCPU)、または特定用途向け集積回路 (Application-Specific Integrated Circuit、略してASIC)、または本発明の本実施形態を実施するように構成された1つ以上の集積回路であってもよい。

【0188】

第1の受信器703は、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声 - テキスト変換要求を受信するように構成される。

【0189】

プロセッサ701は、少なくとも2つの端末によって送信された音声 - テキスト変換要求において、同じ第1の識別子または同じ第2の識別子を有する端末が同じセッションを有するように、少なくとも2つの端末にセッションを割り当てるように構成される。

【0190】

第2の受信器704は、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームを、パケット交換ドメインを用いて受信するように構成される。

【0191】

プロセッサ701は、音声ストリームをテキストに変換するようにさらに構成されてもよい。

【0192】

送信器705は、多者通話において端末にテキストを送信するように構成される。

【0193】

さらに、プロセッサ701は、

第1の端末が、少なくとも2つの端末のうちのいずれかの端末であり、データベースが、少なくとも1つのセッションおよび少なくとも1つのセッションに対応する端末識別子を含むとき、データベースが、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子を含む場合、第1の端末によって送信された音声 - テキスト変換要求の第1の識別子または第2の識別子に対応する、データベース内のセッションを第1の端末に割り当て、

データベースが、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子および第2の識別子を含まず、第2の端末によって送信された音声 - テキスト変換要求の第1の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第1の識別子と同じであり、第2の端末によって送信された音声 - テキスト変換要求の第2の識別子が、第3の端末によって送信された音声 - テキスト変換要求の第2の識別子と同じである場合、第2の端末および第3の端末に同じ新しいセッションを割り当てる

ように特にさらに構成されてもよい。

【0194】

任意選択で、テキストは、少なくとも1つの端末の識別子を含んでもよいし、テキストは、少なくとも1つの端末の識別子およびユーザ名を含んでもよい。

【0195】

さらに、プロセッサ701は、

割り当てられたセッションが新しいセッションである場合、割り当てられたセッションと少なくとも2つの端末の識別子との対応関係を確立し、この対応関係をデータベースに追加するか、または

割り当てられたセッションがデータベース内のセッションである場合、追加された識別子が割り当てられたセッションに対応するように、少なくとも2つの端末の識別子のうちにあり、かつデータベースに含まれない端末識別子をデータベースに追加する

ようにさらに構成されてもよい。

【0196】

さらに、図8に示すように、多者通話において音声をテキストに変換するための装置50は、

第5の端末によって送信された終了メッセージを受信するように構成された第3の受信器706をさらに含んでもよい。

【0197】

さらに、プロセッサ701は、データベースから第5の端末の識別子を削除するようにさらに構成されてもよい。

【0198】

プロセッサ701は、セッションに対応する端末の識別子がデータベースにおいて空である場合、データベースからセッションを削除するようにさらに構成されてもよい。

【0199】

本発明の本実施形態で提供される、多者通話において音声をテキストに変換するための装置50によれば、少なくとも2つの端末によって送信された、第1の識別子および第2の識別子を含む音声 - テキスト変換要求が受信され、少なくとも2つの端末によって送信され

10

20

30

40

50

た音声 - テキスト変換要求において同じ第1の識別子または同じ第2の識別子を有する端末に同じセッションが割り当てられるように、少なくとも2つの端末にセッションが割り当てられ、1回のセッションに対応する1回の多者通話において少なくとも1つの端末によって送信される、サンプリングレートが8KHzよりも大きい音声ストリームが、パケット交換ドメインを用いて受信され、音声ストリームがテキストに変換され、多者通話において端末にテキストが送信される。受信される音声ストリームは、パケット交換ドメインを用いて受信される、サンプリングレートが8KHzよりも大きい音声ストリームであるため、パケット交換ドメインは、高いサンプリングレートの音声ストリームの送信をサポートすることができ、受信した音声ストリームがテキストに変換された後に取得されるテキストの精度は高くなる。さらに、この処理は、手動操作を必要とせず、比較的迅速であり、人間の労力と時間を節約する。結論として、本発明で提供される方法を用いることによって、多者通話において音声から変換されたテキストを正確に取得することは、時間を節約する容易な方法で実施することができ、これにより、多者通話において音声から変換されたテキストを取得することが時間と労力を要するものであるまたは不正確であるという従来技術の問題が解決されることが学ばれ得る。

10

【0200】

簡便かつ簡単な説明のために、前述のシステム、装置、およびユニットの詳細な動作プロセスについて、前述の方法の実施形態の対応するプロセスを参照することができるため、本明細書では詳細について再び説明していないことが当業者によって明確に理解されよう。

20

【0201】

本出願で提供されるいくつかの実施形態において、開示されたシステム、装置、および方法が、他の方法で実施され得ることを理解されたい。例えば、説明した装置の実施形態は一例に過ぎない。例えば、ユニットの分割は、論理的な機能の分割に過ぎず、実際の実施では他の分割であってもよい。例えば、複数のユニットまたはコンポーネントは、別のシステムとして組み合わせられるか、もしくは統合されてもよいし、一部の機能は、無視されるか、もしくは実行されなくてもよい。さらに、表示または議論された相互結合または直接結合または通信接続は、いくつかのインターフェースを介して実施されてもよい。装置またはユニット間の間接結合または通信接続は、電子的なまたは他の形態で実施されてもよい。

30

【0202】

別々の部分として説明されているユニットは、物理的に別々であってもなくてもよく、ユニットとして表示されている部分は、物理的なユニットであってもなくてもよく、1つの位置に配置されても、複数のネットワークユニットに分散されてもよい。ユニットの一部または全部は、実施形態の解決策の目的を達成するために実際の必要性に応じて選択されてもよい。

【0203】

さらに、本発明の実施形態における機能ユニットは、1つの処理ユニットに統合されてもよいし、これらのユニットのそれぞれは、物理的に単独で存在してもよいし、2つ以上のユニットが、1つのユニットに統合される。統合ユニットは、ハードウェアの形態で実施されてもよいし、ソフトウェア機能ユニットに加えてハードウェアの形態で実施されてもよい。

40

【0204】

前述の統合ユニットが、ソフトウェア機能ユニットの形態で実施される場合、統合ユニットは、コンピュータ可読記憶媒体に記憶されてもよい。ソフトウェア機能ユニットは、記憶媒体に記憶され、コンピュータデバイス（パーソナルコンピュータ、サーバ、またはネットワークデバイスであってもよい）に、本発明の実施形態で説明した方法のステップの一部を実行するよう命令するためのいくつかの命令を含む。前述の記憶媒体は、プログラムコードを記憶することができる任意の媒体（USBフラッシュドライブ、リムーバブル・ハード・ディスク、読み出し専用メモリ（Read-Only Memory（ROM））、ランダム・ア

50

クセス・メモリ（Random Access Memory（RAM））、磁気ディスク、または光ディスクなど）を含む。

【0205】

最後に、前述の実施形態は、本発明の技術的解決策を説明するためのものに過ぎず、本発明を限定するためのものではないことに留意されたい。本発明は、前述の実施形態を参照して詳細に説明されているが、当業者であれば、本発明の実施形態の技術的解決策の範囲から逸脱することなく、前述の実施形態で説明した技術的解決策をさらに修正することができる、またはその技術的特徴の同等の交換を行うことができることを理解するはずである。

【符号の説明】

10

【0206】

50 装置

501 第1の受信ユニット

502 割当ユニット

503 第2の受信ユニット

504 変換ユニット

505 送信ユニット

506 追加ユニット

507 第3の受信ユニット

508 削除ユニット

20

701 プロセッサ

702 通信バス

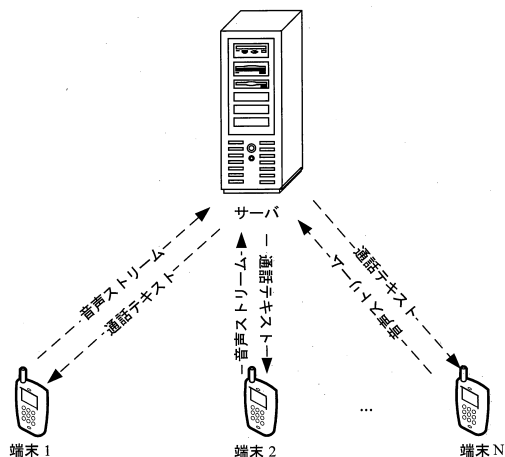
703 第1の受信器

704 第2の受信器

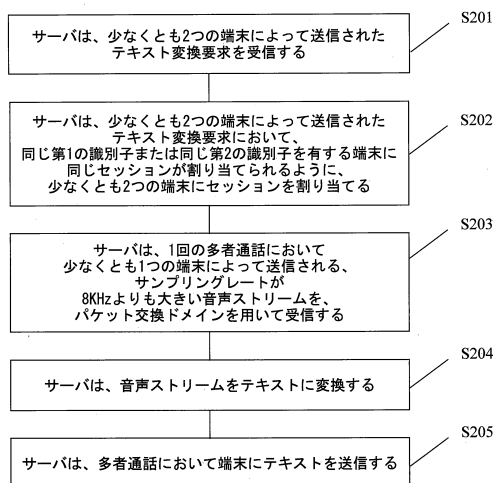
705 送信器

706 第3の受信器

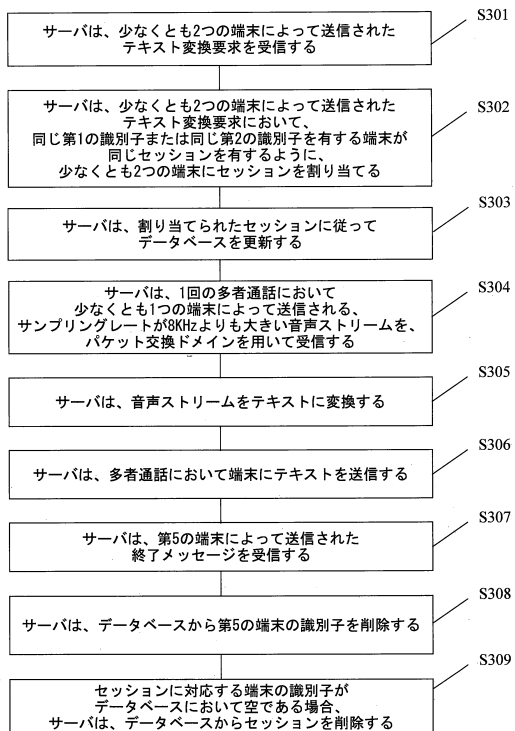
【図1】



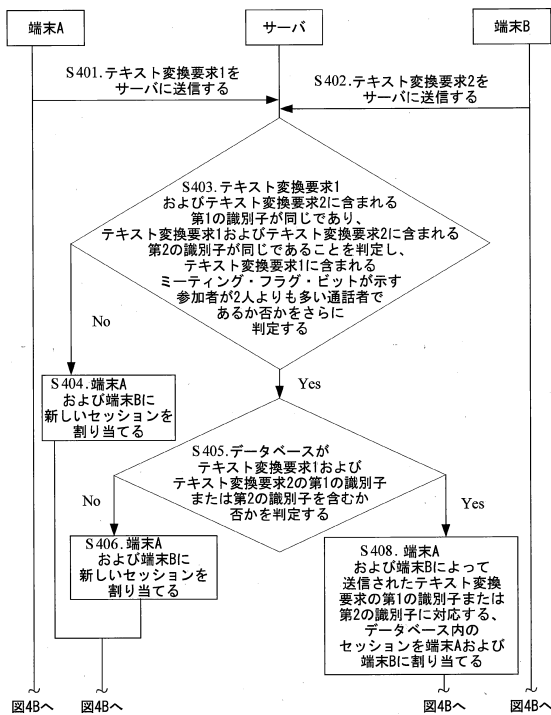
【図2】



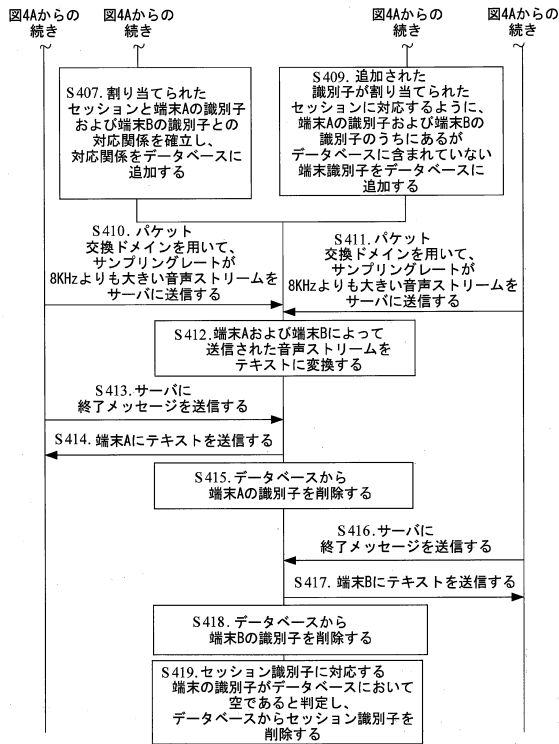
【図3】



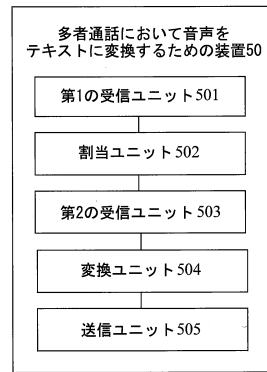
【図4A】



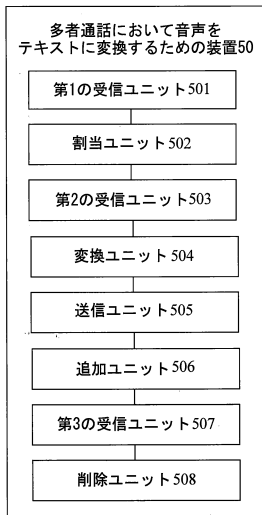
【図4B】



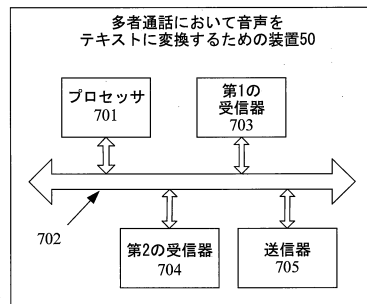
【図5】



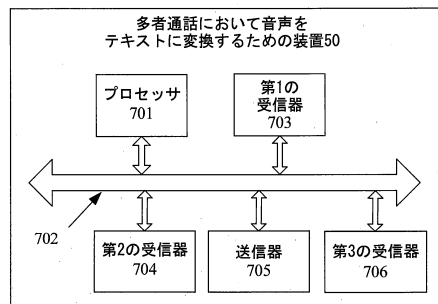
【図6】



【図7】



【図8】



フロントページの続き

(74)代理人 100140534

弁理士 木内 敬二

(72)発明者 王 細 勇

中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為總部 ベン 公楼

(72)発明者 蒋 洪 叡

中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為總部 ベン 公楼

(72)発明者 鄭 偉 軍

中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為總部 ベン 公楼

審査官 石田 紀之

(56)参考文献 特開2013-198066(JP,A)

特開2005-012484(JP,A)

特表2005-513542(JP,A)

米国特許第07236580(US,B1)

(58)調査した分野(Int.Cl., DB名)

H04M 3/56

H04M 3/00

H04M 3/42