(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6417652号
(P6417652)
(45)発行日　平成30年11月7日(2018.11.7)　(24)登録日　平成30年10月19日(2018.10.19)

(51)Int.Cl.　F I
HO4M 3/56 (2006.01)　HO4M 3/56 B

請求項の数 15 (全 23 頁)

(21)出願番号　特願2013-204509(P2013-204509)
(22)出願日　平成25年9月30日(2013.9.30)
(65)公開番号　特開2015-70515(P2015-70515A)
(43)公開日　平成27年4月13日(2015.4.13)
審査請求日　平成28年9月28日(2016.9.28)

(73)特許権者　390002761
キヤノンマーケティングジャパン株式会社
東京都港区港南２丁目１６番６号
(73)特許権者　592135203
キヤノンＩＴソリューションズ株式会社
東京都品川区東品川２丁目４番１１号
(74)代理人　100189751
弁理士　木村　友輔
(74)代理人　100208904
弁理士　伊藤　秀起
(74)代理人　100186059
弁理士　山田　紘子
(72)発明者　矢島　久士
東京都品川区東品川２丁目４番１１号　キヤノンソフトウェア株式会社内
最終頁に続く

(54)【発明の名称】情報処理装置、情報処理システム、情報処理装置の制御方法、情報処理システムの制御方法、およびプログラム

(57)【特許請求の範囲】
【請求項１】
入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置であって、
前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成手段と、
前記合成データ生成手段により生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信する送信手段と、
前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定手段と、
を備え、
前記合成データ生成手段は、前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には、前記合成前音声データ群において前記特定手段により特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成することを特徴とする情報処理装置。
【請求項２】

前記合成データ生成手段は、前記合成データを送信する送信先のクライアント端末ごとに、前記合成前音声データ群から音声データを取得して、送信先のクライアント端末用の前記合成データを生成し、
前記送信手段は、送信先のクライアント端末用に生成された合成データを対象のクライアント端末に送信し、
前記特定手段は、前記合成前音声データ群から、前記送信手段による前記合成データの送信による遅延がある場合には合成に用いない音声データを特定し、
前記合成データ生成手段は、前記特定手段により特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して前記クライアント端末用の合成データを生成することを特徴とする請求項１に記載の情報処理装置。

【請求項３】
前記合成に用いない音声データとは、音声が含まれていないとみなされる無音データであることを特徴とする請求項１又は２のいずれか１項に記載の情報処理装置。

【請求項４】
前記音声が含まれていないとみなされる無音データとは、無音であることを示す付加情報が付加された音声データであることを特徴とする請求項３に記載の情報処理装置。

【請求項５】
入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置の制御方法であって、
前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成ステップと、
前記合成データ生成ステップにより生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信する送信ステップと、
前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定ステップと、
を含み、
前記合成データ生成ステップは、前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には、前記合成前音声データ群において前記特定ステップにより特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成することを特徴とする制御方法。

【請求項６】
入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置で実行可能なプログラムであって、
前記情報処理装置を、
前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成手段と、
前記合成データ生成手段により生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信する送信手段と、
前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定手段として機能させ、
前記合成データ生成手段を、前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には、前記合成前音声データ群において前記特定手段

により特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成する手段として機能させるためのプログラム。

【請求項７】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置とを含む情報処理システムであって、

前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成手段と、

前記合成データ生成手段により生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信する送信手段と、

前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定手段と、

を備え、

前記合成データ生成手段は、前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には、前記合成前音声データ群において前記特定手段により特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成することを特徴とする情報処理システム。

【請求項８】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置とを含む情報処理システムの制御方法であって、

前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成ステップと、

前記合成データ生成ステップにより生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信する送信ステップと、

前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定ステップと、

を含み、

前記合成データ生成ステップは、前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には、前記合成前音声データ群において前記特定ステップにより特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成することを特徴とする制御方法。

【請求項９】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置とを含む情報処理システムを制御するプログラムであって、

前記情報処理システムを、

前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成手段と、

前記合成データ生成手段により生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信する送信手段と、

前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定手段として機能させ、

前記合成データ生成手段を、前記クライアント端末から前記合成データの送信による遅延がある旨の通知を受け付けた場合には、前記合成前音声データ群において前記特定手段により特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成する手段として機能させるためのプログラム。

【請求項１０】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置であって、

前記音声データ群から取得された所定の単位の音声データを、送信先であるクライアント端末に送信する送信手段と、

前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には送信しない音声データを、前記音声データ群から特定する特定手段と、

を備え、

前記送信手段は、前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には、前記音声データ群から、前記特定手段により特定された音声データをスキップして次の音声データを取得し、前記通知を受け付けたクライアント端末に送信することを特徴とする情報処理装置。

【請求項１１】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置の制御方法であって、

前記音声データ群から取得された所定の単位の音声データを、送信先であるクライアント端末に送信する送信ステップと、

前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には送信しない音声データを、前記音声データ群から特定する特定ステップと、

含み、

前記送信ステップは、前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には、前記音声データ群から、前記特定ステップにより特定された音声データをスキップして次の音声データを取得し、前記通知を受け付けたクライアント端末に送信することを特徴とする制御方法。

【請求項１２】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置を制御するプログラムであって、

前記情報処理装置を、

前記音声データ群から取得された所定の単位の音声データを、送信先であるクライアント端末に送信する送信手段と、

前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には送信しない音声データを、前記音声データ群から特定する特定手段として機能させ、

前記送信手段を、前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には、前記音声データ群から、前記特定手段により特定された音声データをスキップして次の音声データを取得し、前記通知を受け付けたクライアント端末に送信する手段として機能させるためのプログラム。

【請求項１３】

入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と、前

記クライアント端末から受信した所定の単位の音声データを音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置とを含む情報処理システムであって、
前記音声データ群から取得された所定の単位の音声データを、送信先であるクライアント端末に送信する送信手段と、
前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には送信しない音声データを、前記音声データ群から特定する特定手段と、
を備え、
前記送信手段は、前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には、前記音声データ群から、前記特定手段により特定された音声データをスキップして次の音声データを取得し、前記通知を受け付けたクライアント端末に送信することを特徴とする情報処理システム。
【請求項１４】
入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と、前記クライアント端末から受信した所定の単位の音声データを音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置とを含む情報処理システムの制御方法であって、
前記音声データ群から取得された所定の単位の音声データを、送信先であるクライアント端末に送信する送信ステップと、
前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には送信しない音声データを、前記音声データ群から特定する特定ステップと、
を含み、
前記送信ステップは、前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には、前記音声データ群から、前記特定ステップにより特定された音声データをスキップして次の音声データを取得し、前記通知を受け付けたクライアント端末に送信することを特徴とする制御方法。
【請求項１５】
入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と、前記クライアント端末から受信した所定の単位の音声データを音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置とを含む情報処理システムを制御するプログラムであって、
前記情報処理システムを、
前記音声データ群から取得された所定の単位の音声データを、送信先であるクライアント端末に送信する送信手段と、
前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には送信しない音声データを、前記音声データ群から特定する特定手段として機能させ、
前記送信手段を、前記クライアント端末から前記音声データの送信による遅延がある旨の通知を受け付けた場合には、前記音声データ群から、前記特定手段により特定された音声データをスキップして次の音声データを取得し、前記通知を受け付けたクライアント端末に送信する手段として機能させるためのプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
ネットワークを介して音声を送受信するアプリケーションにおいて、ネットワークの遅延などによる音声の遅延を回復する技術に関する。
【背景技術】
【０００２】
ネットワークを介して音声を送受信するアプリケーションにおいて、ネットワークの揺らぎなどを考慮し、例えば、サーバ上に各クライアントからの上り音声を蓄積するためのキューを用意する場合がある。しかし、キューを持つことにより、キューにデータが蓄積

していくことで遅延が増大する状況も発生する。
【０００３】
クライアントが受信したそれぞれのクライアントの音声から無音を破棄することで遅延を回復するここことが可能であるが、サーバでミキシングする場合、クライアントが受信する下り音声はサーバでミキシングされた１つの音声データとなる。その場合、いずれか１つのクライアントが発話中であれば、そのミキシングした音声は無音ではないため破棄されず、遅延を回復することはできない。
【０００４】
そのため、下り音声をサーバでミキシングしてクライアントに送信する場合には、当該クライアントとサーバ間の下りの通信が揺らぐなどした際に、他のクライアントにはすでに送信済みである音声が、当該クライアントにだけ遅れて送信されるといったことが発生し、その後に受信する音声がすべて遅延する状態を維持してしまうことになる。
【０００５】
以上のように遅延が増大すると具体的には以下のような問題が発生する。複数拠点で遠隔会議をしているとした場合、時間の経過とともに音声の遅延が蓄積し、会議の進行に支障が出る場合がある。例えば、相手が発話していないと思い、発話した場合に音声が重複し、聞き取れない可能性がある。また、質問に対する回答が来ていないと思い再度質問するなど、会話が成り立たない可能性がある。
【０００６】
特許文献１における通信システムは、音声データを送信装置から受信装置に送信するシステム構成であって、送信側装置は、音声データの音量レベルに基づいて無音状態情報を生成して、送信装置は音声データを蓄積して遅延時間を調整するジッタ吸収バッファと、ジッタ吸収バッファに蓄積されている音声データ量が設定した許容蓄積量を超過した場合に、当該許容蓄積量を増加させ、ジッタ吸収バッファに蓄積されている音声データ量が一定時間、設定した許容蓄積量内である場合に、無音状態を示す音声データを廃棄して、当該許容蓄積量をデフォルト値に戻すジッタバッファ調整部を備えたものである。
【先行技術文献】
【特許文献】
【０００７】
【特許文献１】特開２０１２－１２４６８９
【発明の概要】
【発明が解決しようとする課題】
【０００８】
しかしながら、特許文献１に記載の技術においては、あくまで送信端末、受信端末が一対一の関係にある場合に限られており、例えば、会議システムのように、一対多の関係であって、また、１つのクライアントから他のクライアントに直接音声データを送るわけではなく、例えば会議サーバにおいて他のクライアント端末の音声データをミキシングしてから、クライアント端末に音声データを送信する場合には対応できない。
【０００９】
本発明の目的は、音声データを送信する仕組みにおいて、送信する対象の音声データを変えて遅延を回復することが可能な仕組みを提供することである。
【課題を解決するための手段】
【００１０】
本発明は、入力された音声データを所定の単位で情報処理装置に送信するクライアント端末と通信可能な、前記クライアント端末から受信した所定の単位の音声データを、当該音声データの送信元のクライアント端末ごとに合成前音声データ群として蓄積して記憶する蓄積記憶部を備える情報処理装置であって、前記蓄積記憶部に蓄積された、前記クライアント端末ごとの合成前音声データ群から取得された所定の単位の音声データを用いて合成データを生成する合成データ生成手段と、前記合成データ生成手段により生成された前記合成データを、複数のクライアント端末のうち、送信先であるクライアント端末に送信

する送信手段と、前記クライアント端末から通知を受け付けた場合には合成に用いない音声データを、前記蓄積記憶部に蓄積された前記クライアント端末ごとの合成前音声データ群から特定する特定手段と、を備え、前記合成データ生成手段は、前記クライアント端末から通知を受け付けた場合には、前記合成前音声データ群において前記特定手段により特定された前記合成に用いない音声データをスキップし、合成すべき次の音声データを取得して、前記通知を受け付けたクライアント端末用の合成データを生成することを特徴とする。

【発明の効果】

【００１１】

本発明によれば、音声データを送信する仕組みにおいて、送信する対象の音声データを変えて遅延を回復することが可能な仕組みを提供することができる。

【図面の簡単な説明】

【００１２】

【図１】本発明の実施の形態に係るシステム構成を示す図の一例である。

【図２】本発明の実施の形態に係るハードウェアの構成を示す図の一例である。

【図３】本発明の実施の形態に係るソフトウェアの機能構成を示す図の一例である。

【図４】本発明の実施の形態に係るクライアント端末における再生遅延の発生と回復処理を示すイメージの一例を示す図である。

【図５】本発明の実施の形態に係る会議サーバにおける、音声情報がキューに格納される状況を示すイメージの一例を示す図である。

【図６】本発明の実施の形態に係る会議サーバにおける、クライアント端末からの受信状態に基づく遅延状態の発生と回復処理を示すイメージの一例を示す図である。

【図７】本発明の実施の形態に係る会議サーバにおける、クライアント端末への送信状態に基づく遅延状態の発生と回復処理を示すイメージの一例を示す図である。

【図８】本発明の実施の形態に係るクライアント端末における、再生遅延回復処理のフローチャートの一例である。

【図９】本発明の実施の形態に係る会議サーバにおける、クライアント端末からの受信状態に基づく遅延回復処理のフローチャートの一例である。

【図１０】本発明の実施の形態に係るクライアント端末への送信状態に基づく遅延回復処理のフローチャートの一例である。

【図１１】図７において遅延が発生した場合の遅延回復方法をキューのイメージを用いて説明する図の一例である。

【発明を実施するための形態】

【００１３】

以下、本発明の実施の形態を、図面を参照して詳細に説明する。

【００１４】

図１は、本発明の実施の形態に係るシステム構成を示す図の一例である。クライアント端末１０１（複数）と会議サーバ１０２は、ネットワークを介して接続可能である。

【００１５】

会議システム（図１）においては、会議に参加するユーザのうち、いずれか（便宜上、主催者とする）が、クライアント端末１０１から会議サーバ１０２にアクセスして、会議室を予約する。会議室とは、可能の会議スペースであって、後述するところの「招待された参加者」だけが入室可能にすることも可能である。また、フリースペースとして、不特定のユーザが参加できたり、発言はできないが、視聴のみ可能であったりしても良い。

【００１６】

主催者は、会議を特定する会議ＩＤ（あるいは会議室ＩＤ、部屋番号など）、その会議室を使用する時刻などを決定し、特定の参加者を招待する場合には、会議サーバに登録されている参加者の通知先（例えばメールアドレス）などを用いて、参加を呼びかけるようにしてもよい。特定のユーザのみを参加させるためには、前記通知の中に、当該会議室に入室するためのパスワードを記載可能とする。

【００１７】
ここで、各クライアント端末１０１ａ～ｃのユーザ（実際の会議参加者）のマイクロフォン（不図示）、による音声データ、あるいはカメラ（不図示）による撮像データは、図４～図７でそのイメージとして示すように、一旦、会議サーバ１０２に送信される。例えば音声については各々のクライアント端末１０１から会議サーバ１０２に対して、クライアント端末１０１の個数分の上り音声として送信される。
【００１８】
前述の上り音声は、会議サーバ１０２で、複数のクライアント端末１０１から受信した上り音声をミキシングし、他のクライアント端末１０１に配信する（下り音声データ）。ただ、ミキシングに際して、ある上り音声を送信したクライアント端末１０１に送り返される下り音声データには、その端末自身の音声をミキシングすることは不要である。
【００１９】
ここで、クライアント端末１０１と会議サーバ１０２を別筐体として記載しているが、ある１つのクライアント端末１０１が、会議サーバ１０２の機能を同じ筐体に構成されるようにしても良い。
【００２０】
なお、本発明の実施形態においては、会議システムを例として説明を進めるが、必ずしも会議システムに限らず、例えばネットワークを介して音声の交換をし、ネットワークの遅延などによりパケット(音声データを含むデータ)の遅延が発生する場合に利用可能な技術として提供する。
【００２１】
図２は、本発明の実施の形態に係るハードウェアの構成を示す図の一例である。図２に示すように、クライアント端末１０１、会議サーバ１０２は、システムバス２０４を介してＣＰＵ（Ｃｅｎｔｒａｌ　Ｐｒｏｃｅｓｓｉｎｇ　Ｕｎｉｔ）２０１、ＲＡＭ（Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）２０２、ＲＯＭ（Ｒｅａｄ　Ｏｎｌｙ　Ｍｅｍｏｒｙ）２０３、入力コントローラ２０５、ビデオコントローラ２０６、メモリコントローラ２０７、通信Ｉ／Ｆコントローラ２０８等が接続された構成を採る。　ＣＰＵ２０１は、システムバス２０４に接続される各デバイスやコントローラを統括的に制御する。
【００２２】
また、ＲＯＭ２０３あるいは外部メモリ２１１には、ＣＰＵ２０１の制御プログラムであるＢＩＯＳ（Ｂａｓｉｃ　Ｉｎｐｕｔ／Ｏｕｔｐｕｔ　Ｓｙｓｔｅｍ）やＯＳ（Ｏｐｅｒａｔｉｎｇ　Ｓｙｓｔｅｍ）や、各サーバあるいは各ＰＣが実行する機能を実現するために必要な後述する各種プログラム等が記憶されている。また、本発明を実施するために必要な情報が記憶されている。なお外部メモリはデータベースであってもよい。
【００２３】
ＲＡＭ２０２は、ＣＰＵ２０１の主メモリ、ワークエリア等として機能する。ＣＰＵ２０１は、処理の実行に際して必要なプログラム等をＲＯＭ２０３あるいは外部メモリ２１１からＲＡＭ２０２にロードし、ロードしたプログラムを実行することで各種動作を実現する。
【００２４】
また、入力コントローラ２０５は、キーボード（ＫＢ）２０９や不図示のマウス等のポインティングデバイス等からの入力を制御する。
【００２５】
ビデオコントローラ２０６は、ディスプレイ２１０等の表示器への表示を制御する。尚、表示器は液晶ディスプレイ等の表示器でもよい。これらは、必要に応じて管理者が使用する。
【００２６】
メモリコントローラ２０７は、ブートプログラム、各種のアプリケーション、フォントデータ、ユーザファイル、編集ファイル、各種データ等を記憶する外部記憶装置（ハードディスク（ＨＤ））や、フレキシブルディスク（ＦＤ）、あるいは、ＰＣＭＣＩＡ（Ｐｅ

ｒｓｏｎａｌ Ｃｏｍｐｕｔｅｒ Ｍｅｍｏｒｙ Ｃａｒｄ Ｉｎｔｅｒｎａｔｉｏｎａｌ Ａｓｓｏｃｉａｔｉｏｎ）カードスロットにアダプタを介して接続されるコンパクトフラッシュ（登録商標）メモリ等の外部メモリ２１１へのアクセスを制御する。

【００２７】

通信Ｉ／Ｆコントローラ２０８は、ネットワークを介して外部機器と接続・通信し、ネットワークでの通信制御処理を実行する。例えば、ＴＣＰ／ＩＰ（Ｔｒａｎｓｍｉｓｓｉｏｎ Ｃｏｎｔｒｏｌ Ｐｒｏｔｏｃｏｌ／Ｉｎｔｅｒｎｅｔ Ｐｒｏｔｏｃｏｌ）を用いた通信等が可能である。

【００２８】

尚、ＣＰＵ２０１は、例えばＲＡＭ２０２内の表示情報用領域へアウトラインフォントの展開（ラスタライズ）処理を実行することにより、ディスプレイ２１０上に表示することが可能である。また、ＣＰＵ２０１は、ディスプレイ２１０上のマウスカーソル（図示しない）等によるユーザ指示を可能とする。

【００２９】

本発明を実現するための後述する各種プログラムは、外部メモリ２１１に記録されており、必要に応じてＲＡＭ２０２にロードされることによりＣＰＵ２０１によって実行されるものである。さらに、上記プログラムの実行時に用いられる定義ファイルおよび各種情報テーブル等も、外部メモリ２１１に格納されており、これらについての詳細な説明についても後述する。

【００３０】

図３は、本発明の実施の形態に係るソフトウェアの機能構成を示す図の一例である。クライアント端末１０１と会議サーバ１０２のソフトウェア構成要素と各記憶部、およびそれらの間のデータの受け渡し（端末３から音声データ３２８等に関連付けられた点線矢印を除く）を図で示したものである。

【００３１】

まず、音声データの流れについて概略を説明する。クライアント端末１０１が、ユーザから音声の入力を受け付けると、音声データに変換され会議サーバ１０２に送信される。複数のクライアント端末１０１から会議サーバが受信した音声データは、クライアント端末１０１ごとに用意されたキューに格納される。あるクライアント端末１０１に送信される音声データは、当該クライアント端末１０１自体の音声データ（に対応するキュー）を除き、他のクライアント端末１０１に対応するキューから音声データを１つずつ取得してミキシングし、クライアント端末１０１に送信する。ミキシングされた音声データを受信したクライアント端末１０１は、（当該クライアント端末１０１自体の音声データはミキシングされていない）他のクライアント端末１０１のミキシングされた音声データを再生することでユーザは、音声を聴くことが出来る。

【００３２】

クライアント端末１０１は、音声入力部３１１において、ユーザが使用するマイクなどの接続機器により音声データの入力を受け付ける。入力を受け付けた音声データは、音声送信部３１２から、会議サーバ１０２に送信される。ここで、音声データは、一定のサイズ（例えば、時間を基準として１０ミリセカンド分の入力された音声など）に区切られる。以降の説明で「音声データ」と呼ぶ場合には、この一定のサイズに区切られた入力音声のデータを意味することにする。これはあくまで例であり、データのビット数など会議システムでの基準に従うものとする。

【００３３】

クライアント端末１０１から送信された音声データ（上り音声データ）は、会議サーバ１０２の音声受信部３２１において、受信される。受信した音声データは、ミキシング用音声記憶部３２６に格納される。ミキシング用音声記憶部３２６は、会議に参加している各々のクライアント端末１０１に対応してキューを用意し、ある程度の数の音声データを記憶する。図３においては、ミキシング用音声記憶部３２６のうち、各々のクライアント端末１０１に対応してキューの１つ１つの要素を端末１～端末３に対応して縦方向（例で

は７つの矩形）に示している。
【００３４】
音声受信時に、上りチェック部３２２により上り音声データ受信（取得）が原因で、処理の遅延が発生していると判断された場合には、遅延回復処理が行われる。具体的には、この処理は、受信したクライアント端末１０１ごとに行われ、対応するキューに格納されている音声データの数が、サーバ閾値記憶部３２７の上り遅延閾値に基づき、所定数を上回っていると判断された場合に、無音データの削除処理をする。
【００３５】
ここで、各キューに格納される音声データの一単位は１つの矩形に対応するものであり、より詳細には、音声データ３２８と送信済端末情報３２９のペアを格納する。音声データ３２８は、前述のクライアント端末１０１から受信した音声データである。一方、受信された音声データは、他の端末から受信した音声データとミキシングして、会議サーバ１０２から、各々のクライアント端末１０１に送信（ミキシング音声送信部３２３）される。その際、何らかの状態（例えばネットワークの回線速度の違い）により、送信済のクライアント端末１０１と、未送信クライアント端末１０１とが発生する。そこで、後述の遅延回復処理をクライアント端末１０１ごとに処理できるよう、送信済のクライアント端末１０１か、未送信のクライアント端末１０１かの情報を「送信済端末情報３２９」に記載する。データ構成は如何様でもよく、送信済／未送信のクライアント端末１０１が識別できればよい。
【００３６】
なお、あるクライアント端末１０１へ送信する音声データのミキシングに際して、もともと当該クライアント端末１０１から受信した音声データはミキシングする対象から省く。即ち、自分自身以外のクライアント端末１０１から送られた音声データをミキシングする。
【００３７】
下りチェック部３２４は、ミキシング音声送信部３２３がクライアント端末１０１に音声データを送信する際に、ミキシングすべきクライアント端末１０１（即ち、送信するクライアント端末１０１以外のクライアント端末１０１）関する遅延回復処理をする。具体的には、この処理は、ミキシングする際に、サーバ閾値記憶部３２７の下り遅延フラグ(クライアント端末１０１ごとに存在する)が“オン”になっている場合に、各キュー内の無音データをスキップし、次の音声データをミキシングする。
【００３８】
下りチェック部３２４による遅延回復処理の完了後（遅延回復の必要がないと判定された場合を含む）、ミキシングされたデータは、所定のクライアント端末１０１に送信される。なお、会議サーバ１０２の処理において、音声受信部３２１とミキシング音声送信部３２３は、同期した処理である必要はない。
【００３９】
クライアント端末１０１のミキシング音声受信部３１３においては、会議サーバ１０２が、当該クライアント端末１０１以外の音声データをミキシングしたデータを受信し、受信音声記憶部３１６（キュー）に格納する。その際に、端末閾値記憶部３１７の再生遅延閾値に基づき、所定数を上回っていると判断された場合に、無音データの削除処理をする。
【００４０】
音声再生部３１５においては、受信音声記憶部３１６（キュー）からミキシングされた音声を取り出し、再生する。なお、クライアント端末１０１の処理において、ミキシング音声受信部３１３と音声再生部３１５は、同期した処理である必要はない。
【００４１】
音声の遅延回復処理は、クライアント端末１０１、会議サーバ１０２の双方で実行される。それぞれの遅延回復処理の概要を、音声データを格納するキューのイメージを用いて図４（クライアント端末１０１側）、図５～図７（会議サーバ１０２側）について説明す

る。
【００４２】
図４は、本発明の実施の形態に係るクライアント端末における再生遅延の発生と回復処理を示すイメージの一例を示す図である。図４においては、１）遅延が発生していない状態、２）遅延が発生している状態、３）遅延回復方法、について説明する。なお、図４～図７で説明する本発明の実施の形態に係わるキューはすべてＦＩＦＯとして説明する。まず「遅延が発生していない状態」について１）を用いて説明する。
【００４３】
（１）まずクライアント端末１０１は、会議サーバ１０２から音声データ（ミキシングされた音声データ）を受信する。
（２）その音声データを受信音声記憶部３１６に再生する待ち状態のデータとして格納する。
【００４４】
（３）図４の例では、受信した「Ａ１」のみが格納されているが、再生遅延閾値に基づき、所望の範囲の個数であれば遅延とはみなさいとしてよい。キューは、ＦＩＦＯであり、先に格納された音声データが先に取り出され再生される。
【００４５】
（４）キューの先頭にきた音声データは、音声再生部３１５により取り出され再生されることにより、当該クライアント端末１０１のユーザが聞くことが出来る。
【００４６】
以降の図４における説明で、遅延が発生する／しないの違いはあるものの、この（１）～（４）の流れは同じである。
【００４７】
次に「遅延が発生している状態」について２）を用いて説明する。例えば、ネットワークの遅延などにより複数のパケット（便宜上音声データと同じ記号で表す。例ではＡ１～Ａ７の７個）を、ほぼ同時にクライアント端末１０１にて受信したとする。クライアント端末１０１のキュー（受信音声記憶部３１６）には７個の音声データが蓄積され、音声データの再生が遅延することになる。
【００４８】
これ以降の処理において、再生のためキューから削除されていく音声データと、受信してキューに格納する音声データは原則として同一のタイミングであるため、遅延の速度は一定になる（自然に遅延が回復することはない、ということ）。
【００４９】
従って、何らかの理由により、再度の遅延が発生すると、その遅延した音声データの総数は、次第に蓄積されていき、最終的には、会議の通話にユーザが違和感を感ずるレベルに達する。すなわち、カメラで撮影した画像が別に送受信されている場合には、発言の画像と音声が著しくずれる、また他のクライアント端末１０１との発言内容の前後関係がおかしくなってくる、などが考えられる。
【００５０】
従って、前述の通り、再生遅延閾値に基づき、所望の範囲の個数であれば遅延とはみなさいが、所定の範囲を超えた場合には、再生時の遅延回復処理を行う必要がある。
【００５１】
図４の最後に「遅延回復方法」について３）を用いて説明する。すなわち、再生遅延閾値に基づき、所望の範囲の個数を超えたとして、遅延回復処理を行う。キューにある音声データには、Ａ１～Ａ７の音声データが格納されているが、このうち「（）」が付与された音声データ（Ａ２、Ａ４、Ａ５、Ａ６を例えば（Ａ２）と記載したもの）については、「無音データ」であるとする。ここで無音データとは、システム上「音声が含まれていないとみなす音声データ」である。音声データが無音データであるか否かの判定については「特開２０００－３１２２２３」などにより周知の技術であるため、説明を省略する。
【００５２】

音声データの再生には、キューの先頭から音声データを１つずつ取得して再生するが、取得する際にキューに格納されている音声データの数をカウントし、その数が、再生遅延閾値に基づき、所望の範囲の個数を超えた場合には、遅延回復処理を開始する。その場合、キューに格納されているすべての音声データを確認し、無音データを削除していく。なお、遅延回復処理は、他の処理に割り込まれることなく一気に処理を行う。

【００５３】

図５は、本発明の実施の形態に係る会議サーバにおける、音声情報がキューに格納される状況を示すイメージの一例を示す図である。図５は、遅延が発生していない場合の状態を示す。

（１）まず、会議サーバ１０２は、クライアント端末１０１ごとに音声データを受信する。

【００５４】

（２）次に、受信した音声データを、ミキシング用音声記憶部３２６においてクライアント端末１０１ごとに用意したキューに格納する。格納する情報は、図３で説明したように音声データ３２８だけではなく、他のクライアント端末１０１が当該音声データをミキシングして送信済であるかどうかを、他のクライアント端末１０１ごとに記憶する送信済端末情報３２９をも格納する。両者（３２８、３２９）を合わせて、便宜上音声情報と呼ぶことにする。キューに積んだ直後の初期状態では、他のいずれのクライアント端末１０１にも送信していない。なお、この送信済であるか否かの情報は、他のすべてのクライアント端末１０１の配列を用意したフラグであってもよいし、あるいは送信済端末の一覧、あるいは送信済でない端末の一覧など、いずれのデータ構成であってもよい。

【００５５】

（３）次に、各キューの音声データ（キューの先頭の音声情報に含まれる音声データ）を取り出しミキシングする。前述の通り、音声データを送信しようとするクライアント端末１０１に対応するキューを除き、その他のキューの音声データをミキシングする。

（４）ミキシングした音声データをクライアント端末１０１に送信する。

【００５６】

図６は、本発明の実施の形態に係る会議サーバにおける、クライアント端末からの受信状態に基づく遅延状態の発生と回復処理を示すイメージの一例を示す図である。図６では、まず「１）上り音声データ取得（クライアント端末１０１から会議サーバ１０２に送信された音声データ取得）が原因で、遅延が発生する状況」を説明する。

【００５７】

（１）まず、クライアント端末１０１Ａからは、遅延することなく音声データを受信しているとする。一方、クライアント端末１０１Ｂからは、ネットワークの問題などで、７つの音声データ（Ｂ１～Ｂ７）を受信する。

【００５８】

（２）従って、クライアント端末１０１Ａに対応するキューは、１つの音声情報が格納されるが、クライアント端末１０１Ｂでは、７つの音声情報が対応するキューに格納する。

【００５９】

（３）クライアント端末１０１Ｃに送信するためのミキシングは、クライアント端末１０１Ａ用、Ｂ用のキューから取り出した音声データを用いる。クライアント端末１０１Ａ用のキューからは、音声データ「Ａ７」（「Ｂ７」と同じ時間の音声）が取得され用いられるが、クライアント端末１０１Ｂ用のキューからは、音声データ「Ｂ１」（「Ｂ７」より６単位時間前の音声）が遅延データとして残っている。従って、それぞれ時間のずれた「Ａ７＋Ｂ１」というミキシング音声データが生成される。

【００６０】

前述のクライアント端末１０１における再生の場合と同様に、ミキシングのためキューから削除されていく音声情報と、受信してキューに格納する音声情報は原則として同一の個数であるため、遅延の速度は一定になる（自然に遅延が回復することはない、というこ

と）。従って、何らかの理由により、クライアント端末１０１Ｂから受信する音声データのみに対して遅延が発生すると、その遅延した音声データの総数は、次第に蓄積されていく。

【００６１】

（４）最終的にクライアント端末１０１に送信された（ミキシングされた）音声データは、ユーザからみて問題と判断されるレベルに達する。即ち、ミキシングされた音声データを受信したクライアント端末１０１以外のクライアント端末１０１の音声データがミキシングされているわけだが、時間的なずれがあり、会話が成り立たない可能性もある。

【００６２】

次に、前述の遅延に対する「２）遅延回復方法」を説明する。例として１）のようにクライアント端末１０１Ｂから受信したキューにおいて、所定の数の範囲を超える音声情報が蓄積されたとする（上り遅延閾値に基づいて判定）。所定の数の範囲を超える音声情報が蓄積されたと判定されると遅延回復処理が開始される。「２）」の図の左側にあるキューには、Ｂ１～Ｂ７のうち、無音データに対応する音声情報（「（）」が付与されたもの）が４つあるため、これを削除する。無音データかどうかの判定は、クライアント端末１０１で行い、送信する音声データに付加される。会議サーバ１０２においては、実際の音声データ内部を解析するのではなく、クライアント端末１０１から送信された音声データの付加情報により、無音データであるか否かを判定する。ただし、無音データかどうかの判定は会議サーバ１０２で行ってもよい。なお、会議サーバにおいては、音声データの下り処理（クライアント端末１０１への音声データ送信）においても、音声データが無音であるか否かを判定する場合があるが、同様とする。残りは右側のキューの通り、３つ（Ｂ１、Ｂ３、Ｂ７）の音声情報が格納されることになる。これらが、すべてのキュー（クライアント端末１０１ごとに用意されている）ごとに、実行される。なお、遅延回復処理は、他の処理に割り込まれることなく一気に処理を行う。

【００６３】

図７は、本発明の実施の形態に係る会議サーバにおける、クライアント端末への送信状態に基づく遅延状態の発生と回復処理を示すイメージの一例を示す図である。図７では、まず「３）下り音声データ取得（会議サーバ１０２からクライアント端末１０１へのミキシング音声データ送信）が原因で、遅延が発生する状況」を説明する。

（１）会議サーバ１０２は、クライアント端末１０１Ａ、Ｂの双方から音声データを受信する。

【００６４】

（２）各々の対応するキューに、音声情報（音声データと送信済端末情報を関連付けた情報）を格納する。図７の３）では、それぞれ７つの音声情報が格納されている。

【００６５】

（３）クライアント端末１０１ごとの各キューから、音声データを取り出す。各クライアント端末１０１に送るための音声データをミキシングするため、クライアント端末１０１ごとに、自分自身を除く他のクライアント端末１０１のすべての音声データを取り出す。

【００６６】

（４）次にミキシングした音声データを、クライアント端末１０１Ｃに送信する。その際に、ネットワークの状態などにより、遅延が発生することがある。一方、クライアント端末１０１Ｂへ送信する音声(クライアント端末１０１Ａとクライアント端末１０１Ｃの音声データをミキシングした音声データ)は遅延なく送信できたとする。その場合、クライアント端末１０１Ｂには送信済みだが、クライアント端末１０１Ｃには未送信のクライアント端末１０１Ａの音声がキューに存在することになる。音声情報の送信済端末情報３２９に、未送信クライアント端末１０１が１つでもあれば、キューに格納された音声情報を削除することは出来ない。削除されなくとも、クライアント端末１０１Ｂでは、ミキシング／送信で無視されるデータとなるが、クライアント端末１０１Ｃでは無視できないため、クライアント端末１０１Ａで再生される音声と、クライアント端末１０１Ｃで再生さ

れる音声との時間間隔が累積されてきて、徐々に会話の時間的な前後関係が不自然になることも想定される。

次に、図１１をもちいて前述の遅延に対する「４）遅延回復方法」を説明する。

【００６７】

図１１は、図７において遅延が発生した場合の遅延回復方法をキューのイメージを用いて説明する図の一例である。

【００６８】

例としてクライアント端末１０１Ｃに送信する音声で遅延が発生している場合(下り遅延フラグにて判断)、クライアント端末１０１Ａおよびクライアント端末１０１Ｂの音声をキューから取出しミキシングする際に、無音データをスキップし、無音ではない音声データをミキシングすることで、遅延を回復する。なお、スキップした無音データに未送信端末が存在しない場合はキューから削除される。「４）」の図のキューには、Ｂ１～Ｂ７のうち、無音データに対応する音声情報（「（）」が付与されたもの）が４つあるため、これをスキップする。これらが、すべてのキュー（クライアント端末１０１ごとに用意されている）ごとに、実行される。なお、遅延回復処理は、他の処理に割り込まれることなく一気に処理を行う。

【００６９】

図８～図１０のフローチャートは、クライアント端末１０１、または会議サーバ１０２における遅延に対して、遅延回復の処理を説明するものである。

【００７０】

図８は、本発明の実施の形態に係るクライアント端末における、再生遅延回復処理のフローチャートの一例である。図８のフローチャートの各ステップ（Ｓ８０１～Ｓ８１３）はクライアント端末１０１のＣＰＵ２０１により、また、Ｓ８１４～Ｓ８１５は会議サーバ１０２のＣＰＵ２０１により実行される。図８のフローチャートは、図４のキューのイメージに対応する。

Ｓ８０１においては、会議サーバから送信されたミキシング済音声データを受け付ける。

【００７１】

Ｓ８０２においては、クライアント端末１０１の受信音声記憶部３１６のキューに、前記受信したミキシング済音声データを格納する（エンキュー）。

Ｓ８０３においては、前記キューに格納された音声データの数をカウントする。

【００７２】

Ｓ８０４においては、キューに所定の数の範囲を超える音声データが蓄積されているか判定する（再生遅延閾値に基づいて判定）。

【００７３】

Ｓ８０５においては、所定の数の範囲外音声データが蓄積されたか否かの判定に基づき分岐する。所定の数が範囲内である場合は、ＮＯに進み、Ｓ８１０に進む。所定の数が範囲外である場合には、ＹＥＳに進み＜無音破棄＞のルーチンを実行するためＳ８０６に進む。

【００７４】

このＳ８０６～Ｓ８０９の処理は、他の処理に割り込まれることなく一気に処理を行う。ここで、クライアント端末１０１のキューの全ての音声データについてすべてのチェックを行う。

Ｓ８０６においては、キューの先頭の１つの音声データをデキューする（取り出す）。

Ｓ８０７においては、取り出した音声データが、無音データであるか否かをチェックする。

【００７５】

Ｓ８０８においては、チェックした音声データが無音であると判定された場合には、そのまま（音声データをキューに戻すことなく）次の音声データをチェックすべくＳ８０６に戻る。即ち、無音であると判定された音声データは破棄される。また、チェックした音

声データが無音ではないと判定された場合には、Ｓ８０９において、（Ｓ８０６でキューから取り出した）音声データを、再度キューにエンキューする（格納する）。

【００７６】

前述の通り、キューはＦＩＦＯであり、また、Ｓ８０６～Ｓ８０９の処理は、他の処理に割り込まれることなく、全ての音声データに対して一度ずつ処理を行うため、無音データは全て削除され、図４の３）に記載された左のキューから右のキューのように変わった後でも、残った音声データの時間的順序性は確保される。

【００７７】

以上で、受信音声記憶部３１６におけるミキシング音声データを受信したクライアント端末１０１のキューにおいて遅延が発生した場合には、その遅延回復処理を行った。

【００７８】

次にＳ８１０において、再度、キューに所定の数の範囲を超える音声データが蓄積されているかカウントする。

【００７９】

Ｓ８１１においては、所定の数の範囲を超える音声データが蓄積されているか判定する（再生遅延閾値と同じ値でもよいし、異なる値でもよい）。すなわち、遅延回復処理を完了した結果として十分な効果が得られたか否かを判断する。

【００８０】

Ｓ８１２においては、所定の数の範囲を超える音声データが蓄積されている場合（ＹＥＳの場合）には、Ｓ８１３に進む。そうでない場合（ＮＯの場合）には、Ｓ８０１（音声の受信）に戻る。なお、音声の再生は、非同期処理にて実行されおり、遅延回復処理とは無関係であるため、フローチャートとしては図示していない。

【００８１】

Ｓ８１３においては、会議サーバ１０２に、遅延回復処理の通知を行う。即ち、無音データを削除すると言う方式では、クライアント端末１０１においては、これ以上の対応は出来ない。従って、その場合は、会議サーバ側にも、遅延回復のための支援を要請することになる。また、会議サーバ１０２に、遅延回復処理の通知を行った後、Ｓ８０１に戻る。一方、会議サーバ１０２のＳ８１４においては、クライアント端末１０１のＳ８１３からの通知を受け取る。

【００８２】

Ｓ８１５においては、「下り遅延フラグ」をオンにして、クライアント端末１０１から回復支援処理の支援要請があることを示す。これにより、会議サーバ１０２で動作している「下り音声遅延回復」の処理が起動されるようにしても良い。ただし、「下り音声遅延回復」は、クライアント端末１０１からの要請の有無にかかわらず、独自のルーチンにおける判断で起動されても良い。

以上で、クライアント端末１０１における遅延回復処理に関するフローチャートの説明を完了する。

【００８３】

図９は、本発明の実施の形態に係る会議サーバにおける、クライアント端末からの受信状態に基づく遅延回復処理のフローチャートの一例である。図９のフローチャートの各ステップは会議サーバ１０２のＣＰＵ２０１により実行される。下記の処理は、会議サーバ１０２において、クライアント端末１０１ごとに用意された個々のキューごとに処理される。また、図９のフローチャートで説明する処理は、クライアント端末１０１から会議サーバ１０２が受信する音声データなので、上り音声データに関する音声遅延回復処理の説明である。図９のフローチャートは、図５および図６のイメージに対応する。

【００８４】

Ｓ９０１においては、クライアント端末１０１から音声データ（正確には音声データを含む通信用のパケット）を受信する。ここで、フローチャートを１つだけ記載しているが、具体的には、１つの受信部で受け付けて、そのデータから複数あるクライアント端末１０１のいずれから送信されたデータであるかを判別しても良いし、そもそも、会議サーバ

１０２とクライアント端末１０１の接続が確立されると“クライアント端末１０１ごと”に受信部がマルチスレッドで生成されて、特定のクライアント端末１０１の受信部としてもよい。ただし、Ｓ９０２以降では、受信部でクライアント端末１０１が識別されているものとして、識別された１つのクライアント端末１０１に対応する処理とする。

【００８５】

Ｓ９０２においては、Ｓ９０１で受信した音声データをミキシング用音声記憶部３２６の識別された１つのクライアント端末１０１に対応するキューに、既に音声情報（前述のパケット内の音声データと、前述の送信済端末情報３２９）が幾つ格納されているかをチェックする。

【００８６】

Ｓ９０３においては、前記音声情報の数が、サーバ閾値記憶部３２７の上り遅延閾値に従って所定数を上回っているか否かを判断する。

【００８７】

Ｓ９０４においては、キューにおける音声情報の数が、上り遅延閾値に従って所定数を上回っている、と判定された場合（ＹＥＳの場合）には、Ｓ９０５に進む。上回っていない、と判定された場合（ＮＯの場合）には、Ｓ９０９に進む。

【００８８】

Ｓ９０９においては、キューに格納された音声情報の数が、上り遅延閾値に従って所定数を上回っていない、即ち、まだ（遅延回復処理をすることなく）音声情報を格納できると判断されるため、キューに格納（エンキュー）して、次の音声受信のためＳ９０１に戻る。

【００８９】

Ｓ９０５に進んだ場合には、Ｓ９０５からＳ９０８の処理を、キューに格納されている全ての音声情報に対して繰り返す。その間、Ｓ９０５からＳ９０８のループは抜けない。

【００９０】

上り遅延閾値に従って所定数を上回っている、と判定された場合には、Ｓ９０５において、キューから１つの音声情報を抽出（デキュー。キューから取り出すこと）する。

【００９１】

Ｓ９０６においては、Ｓ９０５において抽出した音声情報に含まれる音声データが、“無音データ”であるか否かを判定する。　Ｓ９０７においては、Ｓ９０５において抽出した音声情報に含まれる音声データが“無音データ”である場合（ＹＥＳの場合）、Ｓ９０５に戻る。すなわち、デキューした音声情報は無音データであるので、キューに戻さない（削除する）。

【００９２】

Ｓ９０８においては、デキューした音声情報は無音データではないので、エンキューする（キューの最後に戻す）。

【００９３】

これらのＳ９０５からＳ９０８の処理により、キューの中の無音データの数分だけ音声情報が削減されることにより、音声遅延を削減するという効果を得ることができる。以上で、図９のフローチャートの説明を完了する。

【００９４】

図１０は、本発明の実施の形態に係るクライアント端末への送信状態に基づく遅延回復処理のフローチャートの一例である。図１０のフローチャートの各ステップは会議サーバ１０２のＣＰＵ２０１により実行される。図１０のフローチャートは、会議サーバ１０２のミキシング用音声記憶部３２６における複数のクライアント端末１０１のキューにおける音声情報（に含まれる音声データ）をミキシングして、クライアント端末１０１に送信する処理である。会議サーバ１０２からクライアント端末１０１に送信するため、上り音声データである。図７のキューのイメージに対応する。

【００９５】

あるクライアント端末１０１に送信するミキシング音声データに、その端末自身の音声

データはミキシングされていない。従って図１０のフローチャートで処理するミキシング用音声記憶部３２６のキューは、送信しようとするクライアント端末１０１以外のキューに対応するキュー１つずつに対して処理される。
Ｓ１００１においては、キューから１つの音声情報を抽出する（デキュー）。
Ｓ１００２においては、図８のＳ８１５で“オン”にされる下り遅延フラグをチェックする。
【００９６】
Ｓ１００３においては、下り遅延フラグが“オン”である場合（ＹＥＳの場合）には、Ｓ１００４に進む。すなわち、クライアント端末１０１側から、会議サーバ１０２側での遅延回復処理が要求されているためである。また、“オフ”である場合（ＮＯ）の場合には、Ｓ１００７に進む。
【００９７】
なお、Ｓ１００２において、下り遅延フラグをチェックするのではなく、上り時のチェック（サーバ閾値記憶部３２７における上り遅延閾値）と同様に、閾値を用いても良い。この閾値を、下り遅延閾値とし、キューにおける音声情報の数が、下り音声閾値に基づいて所定の数に達したらＳ１００３においてＳ１００４（ＹＥＳの方向）、達していない場合にはＳ１００７に進むようにしても良い。
【００９８】
Ｓ１００４においては、Ｓ１００１においてデキューした音声情報に含まれる音声データが、無音データであるか否かをチェックする。
【００９９】
Ｓ１００５においては、Ｓ１００４においてチェックした音声データが“無音データ”である場合（ＹＥＳの場合）には、Ｓ１００６に進み、次の音声情報をキューから取り出して、Ｓ１００４に戻り処理を繰り返す。この繰り返しにより、“無音データ”を無視した処理が進んでいく。
【０１００】
一方、Ｓ１００４においてチェックした音声データが“無音データ”ではない場合（ＮＯの場合）には、Ｓ１００７に進み、前述の通りミキシング音声として、Ｓ１００８によりクライアント端末１０１に送信する。Ｓ１００７においては、ミキシングに使用した音声データは、不要なデータとしてキューに戻していない。しかし、実際には、他のクライアント端末１０１で使用する必要があるかも知れないため、“全てのクライアント端末（この音声データを送信してきたクライアント端末を除く）で、この音声データを送信済みではない”場合には、改めてエンキューする必要がある。　以上で、図１０のフローチャートの説明を完了する。
【０１０１】
なお、上述した各種データの構成及びその内容はこれに限定されるものではなく、用途や目的に応じて、様々な構成や内容で構成されることは言うまでもない。
【０１０２】
以上、一実施形態について示したが、本発明は、例えば、システム、装置、方法、プログラムもしくは記録媒体等としての実施態様をとることが可能であり、具体的には、複数の機器から構成されるシステムに適用しても良いし、また、一つの機器からなる装置に適用しても良い。
【０１０３】
また、本発明におけるプログラムは、図８～図１０のフローチャートの処理方法をコンピュータが実行可能なプログラムであり、本発明の記憶媒体は図８～図１０のフローチャートの処理方法をコンピュータが実行可能なプログラムが記憶されている。なお、本発明におけるプログラムは図８～図１０のフローチャートの各装置の処理方法ごとのプログラムであってもよい。
【０１０４】
以上のように、前述した実施形態の機能を実現するプログラムを記録した記録媒体を、

システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはＣＰＵやＭＰＵ）が記録媒体に格納されたプログラムを読出し実行することによっても、本発明の目的が達成されることは言うまでもない。

【０１０５】

この場合、記録媒体から読み出されたプログラム自体が本発明の新規な機能を実現することになり、そのプログラムを記憶した記録媒体は本発明を構成することになる。

【０１０６】

コンピュータプログラムを供給するための記録媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、ＣＤ－ＲＯＭ、ＣＤ－Ｒ、ＤＶＤ－ＲＯＭ、磁気テープ、不揮発性のメモリカード、ＲＯＭ、ＥＥＰＲＯＭ、シリコンディスク、ソリッドステートドライブ等を用いることができる。

【０１０７】

また、コンピュータが読み出したプログラムを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムの指示に基づき、コンピュータ上で稼働しているＯＳ（オペレーティングシステム）等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【０１０８】

さらに、記録媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるＣＰＵ等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【０１０９】

また、本発明は、複数の機器から構成されるシステムに適用しても、１つの機器からなる装置に適用してもよい。また、本発明は、システムあるいは装置にプログラムを供給することによって達成される場合にも適応できることは言うまでもない。この場合、本発明を達成するためのプログラムを格納した記録媒体を該システムあるいは装置に読み出すことによって、そのシステムあるいは装置が、本発明の効果を享受することが可能となる。さらに、本発明を達成するためのプログラムをネットワーク上のサーバ、データベース等から通信プログラムによりダウンロードして読み出すことによって、そのシステムあるいは装置が、本発明の効果を享受することが可能となる。

なお、上述した各実施形態およびその変形例を組み合わせた構成も全て本発明に含まれるものである。

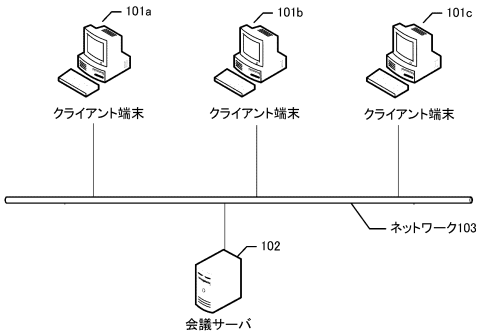
【符号の説明】

【０１１０】

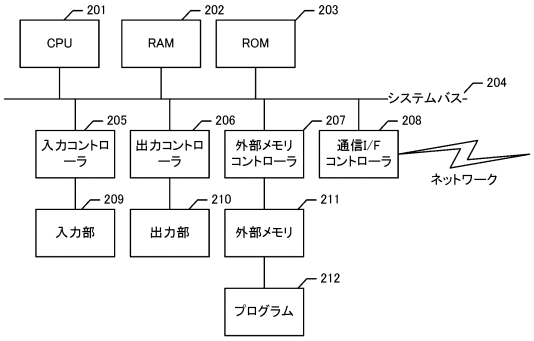
１０１　クライアント端末
１０２　会議サーバ
１０３　ネットワーク
３１１　音声入力部
３１２　音声送信部
３１３　ミキシング音声受信部
３１４　受信音声チェック部
３１５　音声再生部
３１６　受信音声記憶部
３１７　端末閾値記憶部
３２１　音声受信部
３２２　上りチェック部
３２３　ミキシング音声送信部
３２４　下りチェック部

３２５　ミキシング部
３２６　ミキシング用音声記憶部
３２７　サーバ閾値記憶部
３２８　音声データ
３２９　送信済端末情報

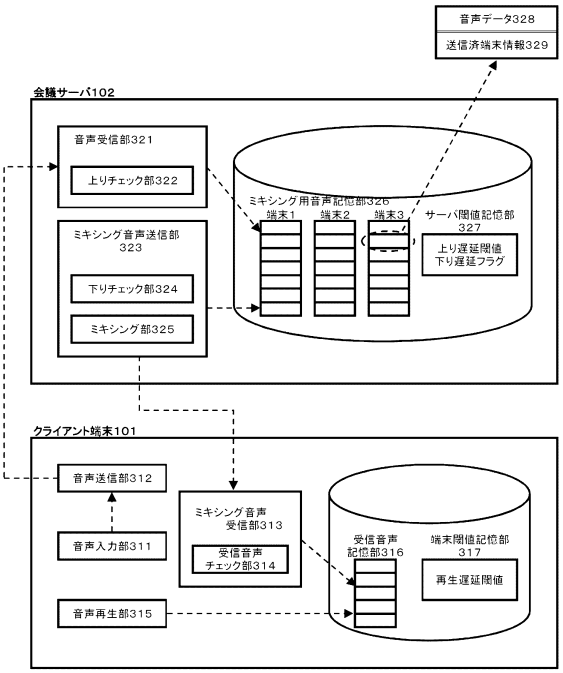
【図１】



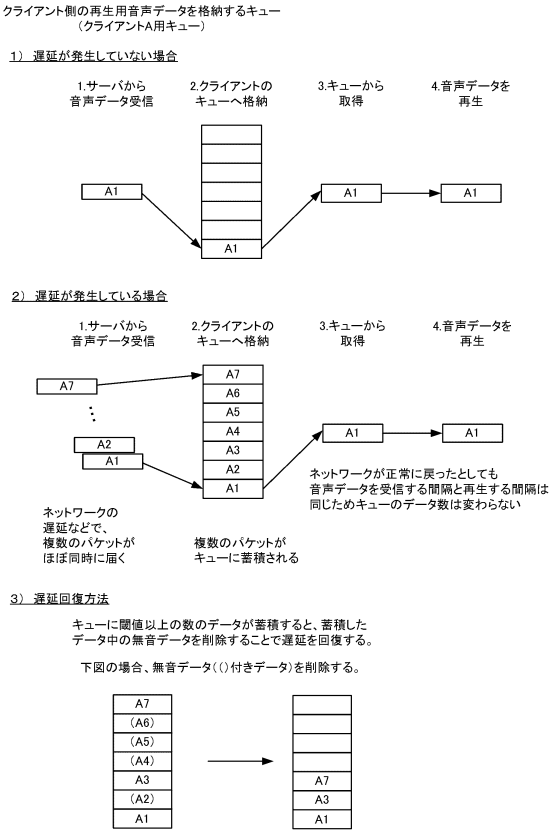
【図２】



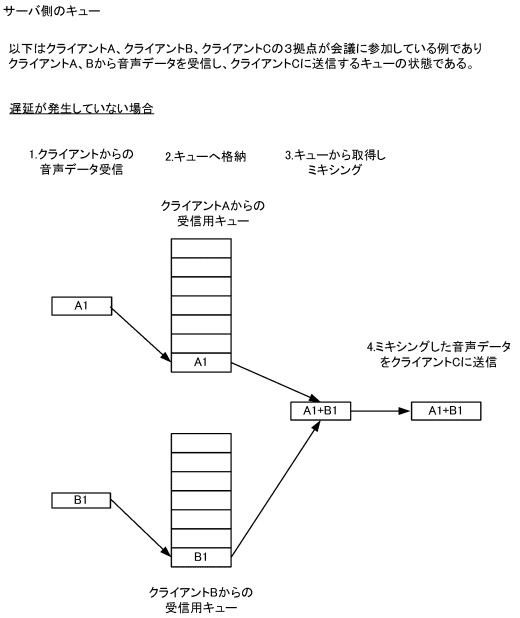
【図３】


音声データ328
送信済端末情報329
会議サーバ102
音声受信部321
上りチェック部322
ミキシング音声送信部
323
下りチェック部324
ミキシング部325
ミキシング用音声記憶部326
端末1
端末2
端末3
サーバ閾値記憶部
327
上り遅延閾値
下り遅延フラグ
クライアント端末101
音声送信部312
音声入力部311
ミキシング音声
受信部313
受信音声
チェック部314
音声再生部315
受信音声
記憶部316
端末閾値記憶部
317
再生遅延閾値

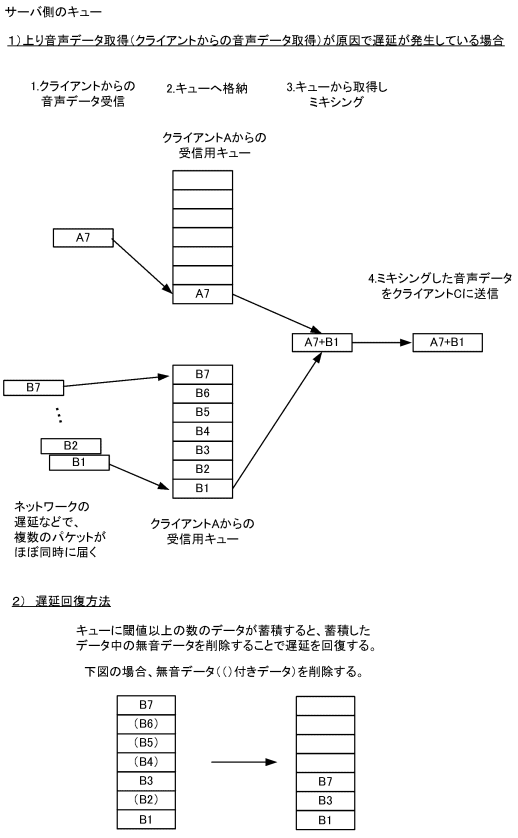
【図４】


クライアント側の再生用音声データを格納するキュー
（クライアントA用キュー）
1）遅延が発生していない場合
1.サーバから
音声データ受信
2.クライアントの
キューへ格納
3.キューから
取得
4.音声データを
再生
A1
A1
A1
A1
2）遅延が発生している場合
1.サーバから
音声データ受信
2.クライアントの
キューへ格納
3.キューから
取得
4.音声データを
再生
A7
A2
A1
A7
A6
A5
A4
A3
A2
A1
A1
A1
ネットワークが正常に戻ったとしても
音声データを受信する間隔と再生する間隔は
同じためキューのデータ数は変わらない
ネットワークの
遅延などで、
複数のパケットが
ほぼ同時に届く
複数のパケットが
キューに蓄積される
3）遅延回復方法
キューに閾値以上の数のデータが蓄積すると、蓄積した
データ中の無音データを削除することで遅延を回復する。
下図の場合、無音データ（()付きデータ）を削除する。
A7
(A6)
(A5)
(A4)
A3
(A2)
A1
A7
A3
A1

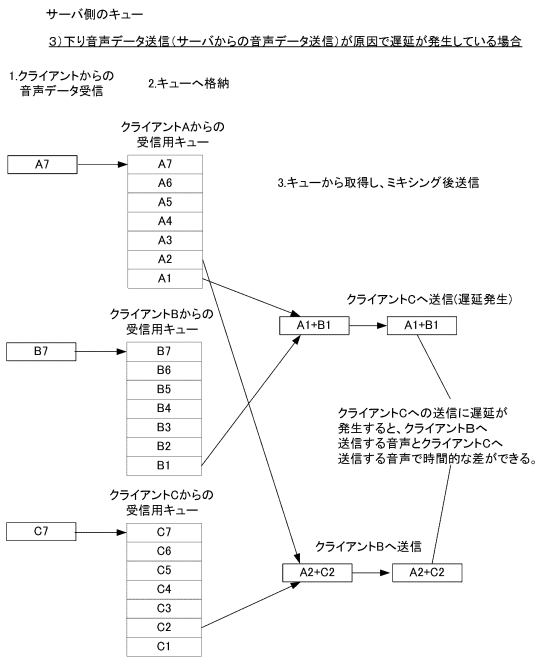
【図５】


サーバ側のキュー
以下はクライアントA、クライアントB、クライアントCの3拠点が会議に参加している例であり
クライアントA、Bから音声データを受信し、クライアントCに送信するキューの状態である。
遅延が発生していない場合
1.クライアントからの
音声データ受信
2.キューへ格納
3.キューから取得し
ミキシング
クライアントAからの
受信用キュー
A1
A1
4.ミキシングした音声データ
をクライアントCに送信
A1+B1
A1+B1
B1
B1
クライアントBからの
受信用キュー

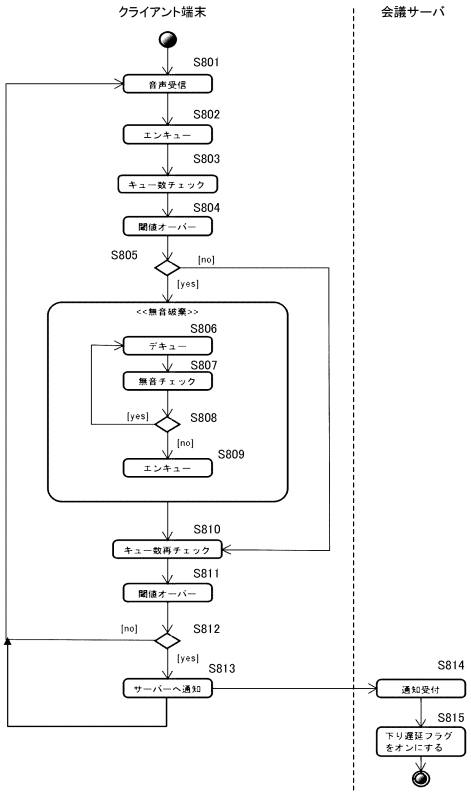
【図６】


サーバ側のキュー
1）上り音声データ取得（クライアントからの音声データ取得）が原因で遅延が発生している場合
1.クライアントからの
音声データ受信
2.キューへ格納
3.キューから取得し
ミキシング
クライアントAからの
受信用キュー
A7
A7
4.ミキシングした音声データ
をクライアントCに送信
A7+B1
A7+B1
B7
B2
B1
B7
B6
B5
B4
B3
B2
B1
ネットワークの
遅延などで、
複数のパケットが
ほぼ同時に届く
クライアントAからの
受信用キュー
2）遅延回復方法
キューに閾値以上の数のデータが蓄積すると、蓄積した
データ中の無音データを削除することで遅延を回復する。
下図の場合、無音データ（()付きデータ）を削除する。
B7
(B6)
(B5)
(B4)
B3
(B2)
B1
B7
B3
B1

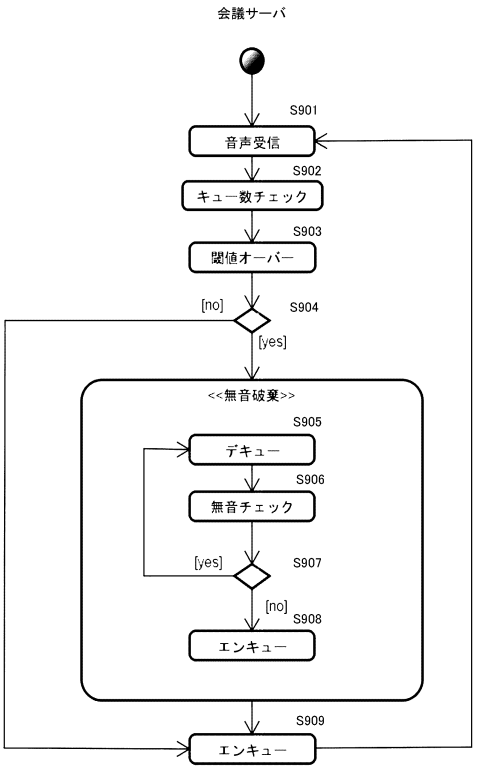
【図７】


サーバ側のキュー
3）下り音声データ送信（サーバからの音声データ送信）が原因で遅延が発生している場合
1.クライアントからの音声データ受信
2.キューへ格納
クライアントAからの受信用キュー
A7
A6
A5
A4
A3
A2
A1
3.キューから取得し、ミキシング後送信
クライアントCへ送信（遅延発生）
A1+B1
A1+B1
クライアントBからの受信用キュー
B7
B6
B5
B4
B3
B2
B1
クライアントCへの送信に遅延が発生すると、クライアントBへ送信する音声とクライアントCへ送信する音声で時間的な差ができる。
クライアントCからの受信用キュー
C7
C6
C5
C4
C3
C2
C1
クライアントBへ送信
A2+C2
A2+C2

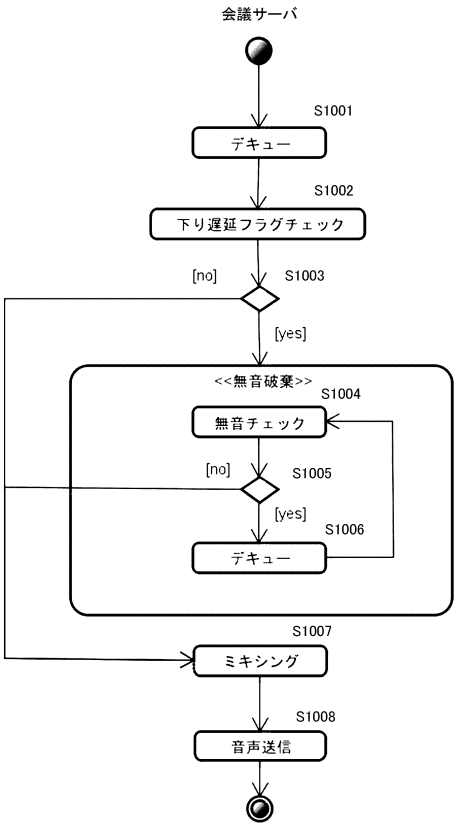
【図８】


クライアント端末
会議サーバ
S801
音声受信
S802
エンキュー
S803
キュー数チェック
S804
閾値オーバー
S805
[no]
[yes]
<<無音破棄>>
S806
デキュー
S807
無音チェック
[yes]
S808
[no]
S809
エンキュー
S810
キュー数再チェック
S811
閾値オーバー
[no]
S812
[yes]
S813
サーバーへ通知
S814
通知受付
S815
下り遅延フラグをオンにする

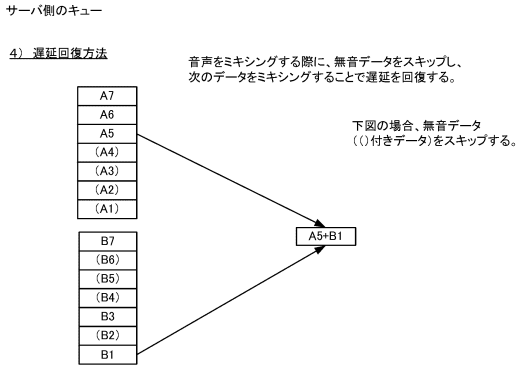
【図９】


会議サーバ
S901
音声受信
S902
キュー数チェック
S903
閾値オーバー
[no]
S904
[yes]
<<無音破棄>>
S905
デキュー
S906
無音チェック
[yes]
S907
[no]
S908
エンキュー
S909
エンキュー

【図１０】


会議サーバ
S1001
デキュー
S1002
下り遅延フラグチェック
[no]
S1003
[yes]
<<無音破棄>>
S1004
無音チェック
[no]
S1005
[yes]
S1006
デキュー
S1007
ミキシング
S1008
音声送信

【図１１】


サーバ側のキュー
4） 遅延回復方法
音声をミキシングする際に、無音データをスキップし、
次のデータをミキシングすることで遅延を回復する。
A7
A6
A5
(A4)
(A3)
(A2)
(A1)
B7
(B6)
(B5)
(B4)
B3
(B2)
B1
下図の場合、無音データ
(()付きデータ)をスキップする。
A5+B1

フロントページの続き

審査官　西巻　正臣

(56)参考文献　特開２００８－０９９３１４（ＪＰ，Ａ）
特開２０１２－１２４６８９（ＪＰ，Ａ）
特開２００２－０７７２３３（ＪＰ，Ａ）
特開２００８－３０６５４１（ＪＰ，Ａ）
特開平１１－２６１６３７（ＪＰ，Ａ）
特開２００３－０２３４９９（ＪＰ，Ａ）
米国特許第０６９４０８２６（ＵＳ，Ｂ１）

(58)調査した分野(Int.Cl.，ＤＢ名)

Ｈ０４Ｂ　７／２４－　７／２６
Ｈ０４Ｌ　１２／００－１２／２６、
１２／５０－１２／９５５
Ｈ０４Ｍ　１／００、
１／２４－　３／００、
３／１６－　３／２０、
３／３８－　３／５８、
７／００－　７／１６、
１１／００－１１／１０、
９９／００
Ｈ０４Ｎ　７／１０、
７／１４－　７／１７３、
７／２０－　７／５６、
２１／００－２１／８５８
Ｈ０４Ｗ　４／００－　８／２４、
８／２６－１６／３２、
２４／００－２８／００、
２８／０２－７２／０２、
７２／０４－７４／０２、
７４／０４－７４／０６、
７４／０８－８４／１０、
８４／１２－８８／０６、
８８／０８－９９／００