

(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2006-254166
(P2006-254166A)

(43) 公開日 平成18年9月21日(2006.9.21)

(51) Int. Cl.		F I		テーマコード (参考)
HO4N	7/15	(2006.01)	HO4N 7/15 630Z	5C164
HO4L	12/18	(2006.01)	HO4L 12/18	5K030
HO4M	3/56	(2006.01)	HO4M 3/56 C	5K201

審査請求 未請求 請求項の数 12 O L (全 30 頁)

(21) 出願番号 特願2005-68917(P2005-68917)
(22) 出願日 平成17年3月11日(2005.3.11)

(71) 出願人 000005108
株式会社日立製作所
東京都千代田区丸の内一丁目6番6号
(74) 代理人 110000198
特許業務法人湘洋内外特許事務所
(72) 発明者 金田 泰
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内
Fターム(参考) 5C164 FA10 VA02S VA07P VA09S VA12P
5K030 HB01 HB02 JT01 JT04 LD08
5K201 AA03 BB09 BD08 CA06 CC04
CC10 EC06 EF04

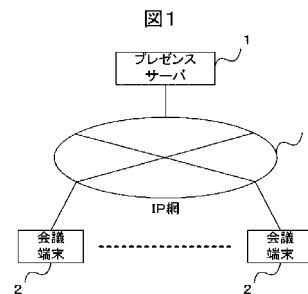
(54) 【発明の名称】 ビデオ会議システム、会議端末および映像サーバ

(57) 【要約】

【課題】ビデオ会議システムにおいて、必要な帯域を低く抑えることができ、且つ、その会議のキーパーソン等の様子を容易に確認できるようにする。

【解決手段】プレゼンスサーバ1は、各会議端末2のユーザの仮想空間における位置情報を管理する。会議端末2は、他の会議端末2の要求に従い自ユーザの映像データを要求元へ送信する。また、仮想空間における自ユーザの位置情報を受付けてプレゼンスサーバ1に送信すると共に、プレゼンスサーバ1から他のユーザの位置情報を受信する。そして、自ユーザおよび他ユーザ各々の位置情報を基に自ユーザの前方に位置する前方ユーザを検出し、前方ユーザ各々の会議端末2から映像データを受信する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

ビデオ会議システムであって、
 複数の会議端末と、前記複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバと、を有し、
 前記複数の会議端末各々は、
 他の会議端末からの映像送信要求に従い、自会議端末のユーザである自ユーザの映像データを要求元の会議端末に送信する映像データ送信手段と、
 前記自ユーザの前記仮想空間上における位置および向きを受け付ける位置受付手段と、
 前記位置受付手段で受け付けた位置および向きを含む自ユーザの仮想位置情報を前記プレゼンスサーバに送信する端末側位置情報送信手段と、
 前記プレゼンスサーバから自会議端末以外の会議端末のユーザである他ユーザ各々の仮想位置情報を受信する端末側位置情報受信手段と、
 前記自ユーザの仮想位置情報および前記他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において前記自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、
 前記前方ユーザ検出手段で検出した前方ユーザ各々の会議端末に映像送信要求を送信して、前記前方ユーザ各々の会議端末からの前方ユーザの映像データを受信する映像データ受信手段と、
 前記映像データ受信手段で受信した前方ユーザの映像データ各々を含む会議画面をディスプレイに表示する表示制御手段と、を有し、
 前記プレゼンスサーバは、
 前記複数の会議端末各々から送られてきたユーザの仮想位置情報を管理する管理手段と、
 前記複数の会議端末の各々に対して、前記管理手段で管理されている当該会議端末以外の各会議端末のユーザの位置情報を当該会議端末に送信するサーバ側位置情報送信手段と、を有すること
 を特徴とするビデオ会議システム。

【請求項 2】

請求項 1 に記載のビデオ会議システムであって、
 前記複数の会議端末各々の表示制御手段は、
 前記前方ユーザ各々の仮想位置情報および前記自ユーザの仮想位置情報により特定される、前記仮想空間上における前記前方ユーザ各々と前記自ユーザとの相対位置に従って、前記前方ユーザ各々の映像データを前記会議画面に複数設けられた表示枠に順番に配置すること
 を特徴とするビデオ会議システム。

【請求項 3】

請求項 1 又は 2 に記載のビデオ会議システムであって、
 前記複数の会議端末各々は、
 前記自ユーザの音声データを、自会議端末以外の会議端末各々に送信する音声データ送信手段と、
 自会議端末以外の会議端末各々から他ユーザの音声データを受信する音声データ受信手段と、
 前記音声データ受信手段で受信した他ユーザの音声データ各々に対して、前記他ユーザ各々の仮想位置情報および前記自ユーザの仮想位置情報により特定される、前記仮想空間における前記他ユーザ各々と前記自ユーザとの相対位置に応じた立体音響処理を施し、立体音響処理された他ユーザの音声データ各々を合成して、立体合成音声データを生成する音声合成手段と、
 前記音声合成手段により生成された立体合成音声データをスピーカから出力する音声制御手段と、

前記自ユーザの発話を検出する発話検出手段と、をさらに有し、
 前記複数の会議端末各々の映像データ送信手段は、
 前記発話検出手段が前記自ユーザの発話を検出した場合に、自会議端末以外の会議端末各々に前記自ユーザの映像データを、発話ユーザの映像データとして送信し、
 前記複数の会議端末各々の映像データ受信手段は、
 前記発話ユーザの映像データを受信し、
 前記複数の会議端末各々の表示制御手段は、
 前記映像データ受信手段で受信した発話ユーザの映像データが前記前方ユーザ以外の他ユーザの映像データである場合に、当該発話ユーザの映像データを前記会議画面に含めてディスプレイに表示すること
 を特徴とするビデオ会議システム。

10

【請求項4】

請求項3に記載のビデオ会議システムであって、
 前記複数の会議端末各々の表示制御手段は、
 前記映像データ受信手段で受信した発話ユーザの映像データが前記前方ユーザ以外の他ユーザの映像データである場合に、前記発話ユーザの映像データを、前記自ユーザの位置情報および前記発話ユーザの位置情報により特定される、前記仮想空間上における前記発話ユーザと前記自ユーザとの相対位置に従って、前記会議画面中の前記前方ユーザ各々の映像データが配置された表示枠の右側あるいは左側にある表示枠に配置すること
 を特徴とするビデオ会議システム。

20

【請求項5】

ビデオ会議システムであって、
 複数の会議端末と、前記複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバと、前記複数の会議端末各々に映像データを送信する映像サーバと、を有し、
 前記複数の会議端末各々は、
 自会議端末のユーザである自ユーザの映像データを前記映像サーバに送信する端末側映像データ送信手段と、
 前記自ユーザの前記仮想空間上における位置および向きを受け付ける位置受付手段と、
 前記位置受付手段で受け付けた位置および向きを含む自ユーザの仮想位置情報を前記プレゼンスサーバに送信する端末側位置情報送信手段と、
 前記映像サーバから会議画面データを受信する端末側映像データ受信手段と、
 前記端末側映像データ受信手段で受信した会議画面データをディスプレイに表示する表示制御手段と、を有し、
 前記映像サーバは、
 前記複数の会議端末各々から当該会議端末のユーザの映像データを受信する映像サーバ側映像データ受信手段と、
 前記プレゼンスサーバから前記複数の会議端末各々の仮想位置情報を受信する映像サーバ側位置情報受信手段と、
 前記複数の会議端末の各々について、当該会議端末の自ユーザの仮想位置情報および他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において当該自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、
 前記複数の会議端末の各々に対して、前記映像サーバ側映像データ受信手段で受信した前記複数の会議端末各々のユーザの映像データのうち、前記前方ユーザ検出手段で検出した当該会議端末の前方ユーザ各々の映像データを含む会議画面データを、当該会議端末へ送信する映像サーバ側映像データ送信手段と、を有し、
 前記プレゼンスサーバは、
 前記複数の会議端末各々から送られてきたユーザの仮想位置情報を受信するプレゼンスサーバ側位置情報受信手段と、
 前記複数の会議端末の各々について、当該会議端末から送られてきたユーザの仮想位置

30

40

50

情報を管理する管理手段と、

前記管理手段で管理されている前記複数の会議端末各々のユーザの仮想位置情報を前記映像サーバに送信するプレゼンスサーバ側位置情報送信手段と、を有すること
を特徴とするビデオ会議システム。

【請求項 6】

請求項 5 に記載のビデオ会議システムであって、

前記映像サーバの映像サーバ側映像データ送信手段は、

前記複数の会議端末の各々に対して、前記前方ユーザ検出手段で検出した当該会議端末の前方ユーザ各々と当該会議端末の自ユーザとの前記仮想空間上における相対位置に従って、当該前方ユーザ各々の映像データの当該会議画面における配置位置を決定し、当該前方ユーザ各々の映像データ各々が、前記決定された配置位置に従って、前記会議画面に複数設けられた表示枠に順番に配置された会議画面データを、当該会議端末に送信すること
を特徴とするビデオ会議システム。

10

【請求項 7】

請求項 5 又は 6 に記載のビデオ会議システムであって、

前記複数の会議端末各々に音声データを送信する音声サーバをさらに有し、

前記複数の会議端末各々は、

自会議端末のユーザである自ユーザの音声データを前記音声サーバに送信する音声データ送信手段と、

前記音声サーバから立体合成音声データを受信する立体合成音声データ受信手段と、

20

前記立体合成音声データ受信手段で受信した立体合成音声データをスピーカから出力する音声制御手段と、

前記自ユーザの発話を検出して、発話検出信号を前記映像サーバに送信する発話検出手段と、をさらに有し、

前記音声サーバは、

前記複数の会議端末各々から当該会議端末のユーザの音声データを受信する音声データ受信手段と、

前記プレゼンスサーバから前記複数の会議端末各々の仮想位置情報を受信する音声サーバ側位置情報受信手段と、

前記複数の会議端末の各々について、前記音声データ受信手段で受信した音声データのうち、当該会議端末以外の会議端末のユーザである他ユーザの音声データ各々に対して、前記他ユーザ各々の仮想位置情報および当該会議端末のユーザである自ユーザの仮想位置情報により特定される、前記仮想空間における前記他ユーザ各々と前記自ユーザとの相対位置に応じた立体音響処理を施し、立体音響処理された他ユーザの音声データ各々を合成して、立体合成音声データを生成する音声合成手段と、

30

前記複数の会議端末各々に対して、前記音声合成手段により生成された当該会議端末に対する立体合成音声データを当該会議端末に送信する立体合成音声データ送信手段と、を有し、

前記映像サーバは、

前記会議端末から発話検出信号を受信する発話検出信号受信手段をさらに有し、

40

前記映像サーバの映像サーバ側映像データ送信手段は、

前記発話検出信号受信手段が発話検出信号を受信した場合に、前記複数の会議端末の各々について、当該発話検出信号の送信元のユーザが前記前方ユーザ検出手段で検出した当該会議端末の前方ユーザ以外のユーザである場合に、当該ユーザの映像データを、発話ユーザの映像データとして会議画面データに含めて、当該会議端末へ送信すること
を特徴とするビデオ会議システム。

【請求項 8】

請求項 7 に記載のビデオ会議システムであって、

前記映像サーバの映像サーバ側映像データ送信手段は、

前記複数の会議端末の各々に対して、当該会議端末に送信する発話ユーザの映像データ

50

の前記会議画面中における配置位置を、前記仮想空間上における前記発話ユーザと当該会議端末の自ユーザとの相対位置に従って決定し、前記発話ユーザの映像データが、前記決定された配置位置に従って、前記会議画面中の前記前方ユーザ各々の映像データが配置された表示枠の右側あるいは左側にある表示枠に配置された会議画面データを、当該会議端末に送信すること

を特徴とするビデオ会議システム。

【請求項 9】

会議端末であって、

他の会議端末からの映像送信要求に従い、自会議端末のユーザである自ユーザの映像データを要求元の会議端末に送信する映像データ送信手段と、

10

前記自ユーザの仮想空間上における位置および向きを受け付ける位置受付手段と、

前記位置受付手段で受け付けた位置および向きを含む自ユーザの仮想位置情報を、ビデオ会議に参加する複数の会議端末各々のユーザの前記仮想空間上における位置を管理するプレゼンスサーバに送信する端末側位置情報送信手段と、

前記プレゼンスサーバから自会議端末以外の会議端末のユーザである他ユーザ各々の仮想位置情報を受信する端末側位置情報受信手段と、

前記自ユーザの仮想位置情報および前記他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において前記自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、

前記前方ユーザ検出手段で検出した前方ユーザ各々の会議端末に映像送信要求を送信して、前記前方ユーザ各々の会議端末からの前方ユーザの映像データを受信する映像データ受信手段と、

20

前記映像データ受信手段で受信した前方ユーザの映像データ各々を含む会議画面をディスプレイに表示する表示制御手段と、を有すること

を特徴とする会議端末。

【請求項 10】

コンピュータ読み取り可能なプログラムであって、

前記プログラムは、コンピュータを、

他の会議端末からの映像送信要求に従い、自会議端末のユーザである自ユーザの映像データを要求元の会議端末に送信する映像データ送信手段と、

30

前記自ユーザの仮想空間上における位置および向きを受け付ける位置受付手段、

前記位置受付手段で受け付けた位置および向きを含む自ユーザの仮想位置情報を、複数の会議端末各々のユーザの前記仮想空間上における位置を管理するプレゼンスサーバに送信する端末側位置情報送信手段と、

前記プレゼンスサーバから自会議端末以外の会議端末のユーザである他ユーザ各々の位置情報を受信する端末側位置情報受信手段と、

前記自ユーザの仮想位置情報および前記他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において前記自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、

前記前方ユーザ検出手段で検出した前方ユーザ各々の会議端末に映像送信要求を送信して、前記前方ユーザ各々の会議端末からの前方ユーザの映像データを受信する映像データ受信手段と、

40

前記映像データ受信手段で受信した前方ユーザの映像データ各々を含む会議画面をディスプレイに表示する表示制御手段と、を有する会議端末として、機能させること

を特徴とするコンピュータ読み取り可能なプログラム。

【請求項 11】

複数の会議端末各々に映像データを送信する映像サーバであって、

前記複数の会議端末各々から当該会議端末のユーザの映像データを受信する映像サーバ側映像データ受信手段と、

前記複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバ

50

バから前記複数の会議端末各々の仮想位置情報を受信する映像サーバ側位置情報受信手段と、

前記複数の会議端末の各々について、当該会議端末の自ユーザの仮想位置情報および他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において当該自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、

前記複数の会議端末の各々に対して、前記映像サーバ側映像データ受信手段で受信した前記複数の会議端末各々のユーザの映像データのうち、前記前方ユーザ検出手段で検出した当該会議端末の前方ユーザ各々の映像データを含む会議画面データを、当該会議端末へ送信する映像サーバ側映像データ送信手段と、を有すること

を特徴とする映像サーバ。

10

【請求項 12】

コンピュータ読み取り可能なプログラムであって、

前記プログラムは、コンピュータを、

前記複数の会議端末各々から当該会議端末のユーザの映像データを受信する映像サーバ側映像データ受信手段と、

前記複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバから前記複数の会議端末各々の仮想位置情報を受信する映像サーバ側位置情報受信手段と、

前記複数の会議端末の各々について、当該会議端末の自ユーザの仮想位置情報および他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において当該自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、

20

前記複数の会議端末の各々に対して、前記映像サーバ側映像データ受信手段で受信した前記複数の会議端末各々のユーザの映像データのうち、前記前方ユーザ検出手段で検出した当該会議端末の前方ユーザ各々の映像データを含む会議画面データを、当該会議端末へ送信する映像サーバ側映像データ送信手段と、を有する映像サーバとして、機能させること

を特徴とするコンピュータ読み取り可能なプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

30

本発明は、立体音響処理された音声を含む会議参加者の映像と共に会議端末から出力することのできるビデオ会議システムの技術に関する。

【背景技術】

【0002】

特許文献 1 には全員表示型ビデオ会議システムが開示されている。この全員表示型ビデオ会議システムでは、各参加者のビデオ会議端末に、会議参加者数分の表示装置を設け、各表示装置に会議参加者の映像を表示して会議参加者全員の映像を常時表示する。

【0003】

また、特許文献 2 にはフロア制御型ビデオ会議システムが開示されている。このフロア制御型ビデオ会議システムでは、発話者を検出し、発話者の映像を選択して各参加者のビデオ会議端末に表示する。

40

【0004】

また、特許文献 3 には 3D オーディオ技術を用いた音声会議システムが開示されている。この音声会議システムでは、音声会議に参加する各参加者の会議端末から、3D オーディオ処理（立体音響処理）された各参加者の音声データを出力する。

【0005】

【特許文献 1】米国特許第 4 4 0 0 7 2 4 号明細書

【特許文献 2】米国特許第 6 3 3 0 0 2 2 号明細書

【特許文献 3】米国特許第 6 3 2 7 5 6 7 号明細書

【発明の開示】

50

【発明が解決しようとする課題】

【0006】

特許文献1に記載の全員表示型ビデオ会議システムでは、参加者一人あたりの映像表示に必要な帯域をBとし、会議参加者数をnとすると、 $(n-1) \times B$ の帯域が必要となる。会議の加者数が多くなると必要な帯域も大きくなり、それに伴いシステムのコストが増加してしまう。

【0007】

一方、特許文献2に記載のフロア制御型ビデオ会議システムでは、参加者一人あたりの映像表示に必要な帯域をBとし、同時発話者数をc ($c < n =$ 会議参加者数)とすると、必要な帯域は $c \times B$ で済む。全員表示型ビデオ会議システムに比べて必要な帯域を小さくできる。しかし、発話者しか表示されないため、その会議のキーパーソン(例えば会議の議長)の様子を確認できない。

10

【0008】

また、特許文献3に記載の3Dオーディオ技術を用いた音声会議システムでは、映像を使用しないため、会議の参加者は当該会議に参加する他の参加者の様子を確認し難い。

【0009】

本発明は上記事情に鑑みてなされたものであり、本発明の目的は、ビデオ会議システムにおいて、必要な帯域を低く抑えることができ、且つ、その会議のキーパーソン等の様子を容易に確認できるようにすることにある。

【課題を解決するための手段】

20

【0010】

上記課題を解決するために、本発明では、ビデオ会議に参加する複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバを設ける。そして、会議端末毎に、プレゼンスサーバを用いて、当該会議端末のユーザが指定した当該ユーザの前記仮想空間における位置および向きに対して前方に位置するユーザ各々を特定し、特定したユーザ各々の映像データを含む会議画面を当該会議端末に表示させる。

【0011】

例えば、本発明のビデオ会議システムは、複数の会議端末と、前記複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバと、を有し、

30

前記複数の会議端末各々は、

他の会議端末からの映像送信要求に従い、自会議端末のユーザである自ユーザの映像データを要求元の会議端末に送信する映像データ送信手段と、

前記自ユーザの前記仮想空間上における位置および向きを受け付ける位置受付手段と、

前記位置受付手段で受け付けた位置および向きを含む自ユーザの仮想位置情報を前記プレゼンスサーバに送信する端末側位置情報送信手段と、

前記プレゼンスサーバから自会議端末以外の会議端末のユーザである他ユーザ各々の仮想位置情報を受信する端末側位置情報受信手段と、

前記自ユーザの仮想位置情報および前記他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において前記自ユーザの前方に位置する少なくともひとつの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、

40

前記前方ユーザ検出手段で検出した前方ユーザ各々の会議端末に映像送信要求を送信して、前記前方ユーザ各々の会議端末からの前方ユーザの映像データを受信する映像データ受信手段と、

前記映像データ受信手段で受信した前方ユーザの映像データ各々を含む会議画面をディスプレイに表示する表示制御手段と、を有し、

前記プレゼンスサーバは、

前記複数の会議端末各々から送られてきたユーザの仮想位置情報を管理する管理手段と

、
前記複数の会議端末の各々に対して、前記管理手段で管理されている当該会議端末以外の各会議端末のユーザの位置情報を当該会議端末に送信するサーバ側位置情報送信手段と

50

、を有する。

【0012】

また、本発明の他のビデオ会議システムは、複数の会議端末と、前記複数の会議端末各々のユーザの仮想空間上における位置を管理するプレゼンスサーバと、前記複数の会議端末各々に映像データを送信する映像サーバと、を有し、

前記複数の会議端末各々は、

自会議端末のユーザである自ユーザの映像データを前記映像サーバに送信する端末側映像データ送信手段と、

前記自ユーザの前記仮想空間上における位置および向きを受け付ける位置受付手段と、

前記位置受付手段で受け付けた位置および向きを含む自ユーザの仮想位置情報を前記プレゼンスサーバに送信する端末側位置情報送信手段と、 10

前記映像サーバから会議画面データを受信する端末側映像データ受信手段と、

前記端末側映像データ受信手段で受信した会議画面データをディスプレイに表示する表示制御手段と、を有し、

前記映像サーバは、

前記複数の会議端末各々から当該会議端末のユーザの映像データを受信する映像サーバ側映像データ受信手段と、

前記プレゼンスサーバから前記複数の会議端末各々の仮想位置情報を受信する映像サーバ側位置情報受信手段と、

前記複数の会議端末の各々について、当該会議端末の自ユーザの仮想位置情報および他ユーザ各々の仮想位置情報に基づいて、前記仮想空間において当該自ユーザの前方に位置する少なくともひとりの他ユーザを、前方ユーザとして検出する前方ユーザ検出手段と、 20

前記複数の会議端末の各々に対して、前記映像サーバ側映像データ受信手段で受信した前記複数の会議端末各々のユーザの映像データのうち、前記前方ユーザ検出手段で検出した当該会議端末の前方ユーザ各々の映像データを含む会議画面データを、当該会議端末へ送信する映像サーバ側映像データ送信手段と、を有し、

前記プレゼンスサーバは、

前記複数の会議端末各々から送られてきたユーザの仮想位置情報を受信するプレゼンスサーバ側位置情報受信手段と、

前記複数の会議端末の各々について、当該会議端末から送られてきたユーザの仮想位置情報を管理する管理手段と、 30

前記管理手段で管理されている前記複数の会議端末各々のユーザの仮想位置情報を前記映像サーバに送信するプレゼンスサーバ側位置情報送信手段と、を有する。

【発明の効果】

【0013】

本発明では、会議端末のユーザが指定した当該ユーザの仮想空間における位置および向きに対して前方に位置するユーザ各々の映像データを当該会議端末に表示するので、ユーザ全員の映像データを会議端末に表示する場合に比べて必要な帯域を小さくできる。また、会議端末のユーザが、仮想空間において所望のユーザが自身の前方に位置するように自身の仮想空間における位置および向きを指定することにより、発話の有無に関わらずに、所望のユーザの様子を確認できる。したがって、本発明によれば、必要な帯域を低く抑えることができ、且つ、その会議のキーパーソン等の様子を容易に確認できる。 40

【発明を実施するための最良の形態】

【0014】

以下に、本発明の実施の形態を説明する。

【0015】

<<第1実施形態>>

図1は本発明の第1実施形態が適用されたビデオ会議システムの概略構成図である。図示するように、本実施形態のビデオ会議システムは、プレゼンスサーバ1および複数の会議端末2がIP(Internet Protocol)網3を介して相互接続されて構成されている。 50

【 0 0 1 6 】

プレゼンスサーバ1は、各会議端末2のユーザの仮想空間上における位置情報を管理する。ここで、仮想空間とは各会議端末2のユーザが会議を行うために仮想的に作り出した空間である。仮想空間の属性には、例えば、空間の大きさ、天井の高さ、壁および天井の反射率・色彩・質感、残響特性、空間内の空気による音の吸収率などがある。

【 0 0 1 7 】

図2はプレゼンスサーバ1の概略構成図である。

【 0 0 1 8 】

図示するように、プレゼンスサーバ1は、IP網3に接続するためのIP網インターフェース部101と、位置情報管理部102と、SIP(Session Initiation Protocol)サーバ処理部103と、位置情報記憶部104と、を有する。

10

【 0 0 1 9 】

図3は位置情報記憶部104の登録内容を模式的に示した図である。図示するように、位置情報記憶部104には、会議端末2のユーザ毎に、レコード1040が記憶されている。レコード1040は、会議端末2のユーザを一意に識別するためのユーザIDを登録するフィールド1041と、当該会議端末2が映像交信に利用するSIP-URI(Uniform Resource Identifier)である映像交信用SIP-URIを登録するフィールド1042と、映像交信用SIP-URIに対応するIPアドレスである映像交信用IPアドレスを登録するフィールド1043と、当該会議端末2が音声交信に利用するSIP-URIである音声交信用SIP-URIを登録するフィールド1044と、音声交信用SIP-URIに対応する音声交信用IPアドレスを登録するフィールド1045と、当該会議端末2のユーザの仮想空間における位置情報である仮想位置情報を登録するフィールド1040と、を有する。ここで、仮想位置情報は、仮想空間におけるユーザの現在位置の座標情報と当該ユーザの視線方向の方位情報とを含む。

20

【 0 0 2 0 】

位置情報管理部102は、位置情報記憶部104に登録されているレコード1040の検索・更新を行う。SIPサーバ処理部103は、位置情報記憶部104に登録されているSIP-URIとIPアドレスとの対応関係を用いて、発側の会議端末2から受信したINVITEメッセージを着側の会議端末2へ送信する。

【 0 0 2 1 】

図4はプレゼンスサーバ1の動作フローを説明する図である。

30

【 0 0 2 2 】

位置情報管理部102は、IP網インターフェース部101を介して会議端末2からユーザIDと共に仮想位置情報を受信すると(S1001)、当該ユーザIDがフィールド1041に登録されているレコード1040を位置情報記憶部104から検索し(S1002)、検索したレコード1040のフィールド1046に登録されている仮想位置情報を該受信した仮想位置情報に更新する(S1003)。

【 0 0 2 3 】

また、位置情報管理部102は、IP網インターフェース部101を介して会議端末2からユーザIDを伴う位置情報送信要求を受信すると(S1004)、位置情報記憶部104から当該ユーザIDがフィールド1041に登録されているレコード1040以外の全てのレコード1040を読み出し(S1005)、該要求の送信元の会議端末2に返信する(S1006)。

40

【 0 0 2 4 】

また、SIPサーバ処理部103は、IP網インターフェース部101を介して会議端末2から、宛先のSIP-URIの指定を伴うINVITEメッセージを受信すると(S1007)、該SIP-URIがフィールド1042あるいは1044に登録されているレコード1040を位置情報記憶部104から検索する(S1008)。そして、該SIP-URIがフィールド1042に登録されているならば、検索したレコード1040のフィールド1043に登録されているIPアドレスを宛先として該INVITEメッセー

50

ジを転送する。一方、該 SIP-URI がフィールド 1044 に登録されているならば、検索したレコード 1040 のフィールド 1045 に登録されている IP アドレスを宛先として該 INVITE メッセージを転送する (S1009)。

【0025】

会議端末 2 は、仮想空間においてユーザが指定した当該ユーザの位置および向き対して前方に位置するユーザ各々を特定し、特定したユーザ各々の映像データを含む会議画面を表示する。図 5 は会議端末 3 の概略構成図である。

【0026】

図示するように、会議端末 2 は、音声入力部 201 と、映像入力部 202 と、音声出力部 203 と、映像出力部 204 と、操作受付部 205 と、オーディオエンコーダ 206 と、ビデオエンコーダ 207 と、オーディオレンダラ 208 と、会議画面生成部 209 と、プレゼンスプロバイダ 210 と、空間モデラ 211 と、IP 網 3 に接続するための IP 網インターフェース部 212 と、RTP (Real-time Transport Protocol) 処理部 213 と、SIP 制御部 214 と、発話検出部 215 と、前方ユーザ検出部 216 と、着席情報作成部 217 と、を有する。

10

【0027】

音声入力部 210 は、マイク 221 で収録した音声信号の入力端子である。映像入力部 202 は、ビデオカメラ 222 で撮像した映像信号の入力端子である。音声出力部 203 は、3D オーディオ対応 (例えば擬似 5.1 チャンネル対応) のヘッドホン (あるいはスピーカ) 223 に接続される音声出力端子である。そして、操作受付部 205 はユーザのポインティングデバイス 225 に対する操作を受け付ける。

20

【0028】

映像出力部 204 は、ディスプレイ 224 に接続される映像出力端子であり、操作受付部 205 に接続されたポインティングデバイス 225 を介してユーザより受け付けた指示に従い、後述する会議画面生成部 209 および着席情報生成部 217 のいずれかを選択し、選択した生成部から出力される映像をディスプレイ 224 に表示する。

【0029】

オーディオエンコーダ 206 は、音声入力部 201 に入力された音声信号をエンコードして音声データを出力する。ビデオエンコーダ 207 は、映像入力部 202 に入力された映像信号をエンコードして映像データを出力する。

30

【0030】

RTP 処理部 213 は、オーディオエンコーダ 206 より出力された音声データを RTP パケットに格納し、該 RTP パケットを IP 網インターフェース部 212 を介して、SIP 処理部 214 より通知された宛先の IP アドレスへ送信する。また、RTP パケット処理部 214 は、ビデオエンコーダ 207 より出力された映像データを RTP パケットに格納し、該 RTP パケットを IP 網インターフェース部 212 を介して、SIP 処理部 214 より通知された宛先の IP アドレスへ送信する。また、RTP 処理部 213 は、IP 網インターフェース部 212 を介して他の会議端末 2 より受信した RTP パケットから音声データを取り出して、該 RTP パケットの送信元アドレスと共にオーディオレンダラ 208 に出力する。さらに、RTP 処理部 213 は、IP 網インターフェース部 212 を介して他の会議端末 2 より受信した RTP パケットから映像データを取り出して、該 RTP パケットの送信元アドレスと共に会議画面生成部 209 に出力する。

40

【0031】

発話検出部 215 は、音声入力部 201 に入力する音声信号からユーザの発話の有無をモニタする。そして、発話 (有音状態) を検出した場合に、発話検出信号を SIP 制御部 214 に出力し、所定期間 (例えば 10 秒) の無音状態を検出した場合に、発話停止信号を SIP 制御部 214 に出力する。

【0032】

空間モデラ 211 は、予め設定されている仮想空間の属性に従い、操作受付部 205 で受け付けた自ユーザのポインティングデバイス 225 に対する操作に応じて当該仮想空間

50

における自ユーザの位置（座標）および視線方向（方位）を決定し、決定した位置および視線方向を含む自ユーザの仮想位置情報を保持すると共に、プレゼンスプロバイダ 210、前方ユーザ検出部 216、着席情報生成部 217 およびオーディオレンダラ 208 に出力する。また、空間モデラ 211 は、プレゼンスプロバイダ 210 から他ユーザ各々の仮想位置情報を受け取って保持すると共に、前方ユーザ検出部 216、着席情報生成部 217 およびオーディオレンダラ 208 に出力する。

【0033】

プレゼンスプロバイダ 210 は、空間モデラ 211 から受け取った自ユーザの仮想位置情報を、IP 網インターフェース部 210 を介してプレゼンスサーバ 1 に定期的に送信する。また、プレゼンスプロバイダ 210 は、IP 網インターフェース部 212 を介してプレゼンスサーバ 1 に位置情報送信要求を定期的に送信し、その応答としてプレゼンスサーバ 1 から他ユーザ各々のレコード 1040 を受信する。そして、受信した他ユーザ各々のレコード 1040 を空間モデラ 211 に通知する。

10

【0034】

前方ユーザ検出部 216 は、空間モデラ 211 から受け取った自ユーザの仮想位置情報および他ユーザ各々のレコード 1040 の仮想位置情報に基づいて、仮想空間において自ユーザの前方に位置する他ユーザを前方ユーザとして検出する。また、検出した前方ユーザ各々の後述する会議画面における配置位置を決定する。そして、検出した前方ユーザ各々を SIP 処理部 214 に通知すると共に、検出した前方ユーザ各々とその配置位置を会議画面生成部 209 に通知する。

20

【0035】

具体的には、図 6 (A) に示すように、まず、前方ユーザ検出部 216 は、仮想空間における自ユーザの位置 2161 および他ユーザ各々の位置 2162 を、それぞれの仮想位置情報に含まれている座標上に決定する。次に、前方ユーザ検出部 216 は、自ユーザの視線方向 2163 を自ユーザの仮想位置情報に含まれている方位に決定する。それから、前方ユーザ検出部 216 は、視線方向 2163 に最も近い前方ユーザの会議画面における配置位置を画面中央とする。また、画面中央とした前方ユーザの右側 n 番目に位置する前方ユーザの会議画面における配置位置を画面中央から右側 n 番目とし、画面中央とした前方ユーザの左側 n 番目に位置する前方ユーザの会議画面における配置位置を画面中央から左側 n 番目とする。この処理を n が予め定められた数（後述する会議画面に設けられた前方ユーザ用の映像表示枠の数に応じて定められた数）となるまで繰り返す。次に、前方ユーザ検出部 216 は、前方ユーザ各々のレコード 1040 に含まれているユーザ ID を SIP 制御部 214 に通知する。また、前方ユーザ毎に、当該前方ユーザのレコード 1040 に含まれているユーザ ID および映像交信用 IP アドレスと当該前方ユーザの会議画面における配置位置とを含む前方ユーザ情報を生成し、会議画面生成部 209 に送信する。

30

【0036】

また、前方ユーザ検出部 216 は、会議画面生成部 217 から受け取った送信元 IP アドレスを映像交信用 IP アドレスとする他ユーザのレコード 1040 を、発話ユーザのレコード 1040 とし、発話ユーザの後述する会議画面における配置位置を決定する。そして、発話ユーザの配置位置を会議画面生成部 209 に通知する。

40

【0037】

具体的には、図 6 (B) に示すように、まず、前方ユーザ検出部 216 は、仮想空間における自ユーザの位置 2161 および発話ユーザ各々の位置 2164 を、それぞれの仮想位置情報に含まれている座標上に決定する。次に、前方ユーザ検出部 216 は、自ユーザの視線方向 2163 を自ユーザの仮想位置情報に含まれている方位に決定する。それから、前方ユーザ検出部 216 は、発話ユーザ各々の会議画面における配置位置を、視線方向 216 に対して自ユーザから右および左のいずれに位置するか否か、および、右および左の何番目に位置するかに従って決定する。例えば、視線方向 2163 の右側 2 番目に位置する発話ユーザの会議画面における配置位置を、前方ユーザの配置位置より右側であって画面中央側から 2 番目とする。そして、発話ユーザ毎に、当該発話ユーザのレコード 10

50

40に含まれているユーザIDおよび映像交信用IPアドレスと当該前方ユーザの会議画面における配置位置とを含む発話ユーザ情報を生成し会議画面生成部209に送信する。

【0038】

会議画面生成部209は、会議画面雛形データを保持している。図7は、会議画面雛形データ2090の表示例を示す図である。この例において、会議画面雛形データ2090は、ユーザ映像表示枠2091およびユーザID表示欄2092の組を複数備えて構成される。複数の組のうち、画面の中央部分に配置される複数(図7では3つ)の組2093は、前方ユーザの映像データおよびユーザIDを表示するための組である。また、複数の組2093の右側および左側に配置される少なくとも1つ(図7では1つ)の組2094は、前方ユーザ以外の発話ユーザの映像データおよびユーザIDを表示するための組である。会議画面生成部209は、映像出力部204を介してディスプレイ224に会議画面雛形データ2090を表示する。また、RTP処理部213から送信元アドレスと共に受け取った各映像データを、前方ユーザの映像データについては複数の組2093のいずれかに当該前方ユーザのユーザIDと共に表示し、前方ユーザ以外の発話ユーザの映像データについては少なくとも1つの組2094のいずれかに当該発話ユーザのユーザIDと共に表示する。

10

【0039】

図8および図9は、会議画面生成部209の動作フローを説明する図である。

【0040】

まず、会議画面生成部209は、前方ユーザ検出部216から受信した前方ユーザ情報の中から会議画面の配置位置が画面中央である前方ユーザ情報を抽出する(S2001)。次に、抽出した前方ユーザ情報の映像交信用IPアドレスを送信元アドレスとする映像データを、RTP処理部213から送信元アドレスと共に受信した映像データの中から検索する(S2002)。そして、S2001で抽出した前方ユーザ情報のユーザIDおよびS2003で検索した映像データの配置位置を、会議画面雛形データ2090中の複数の組2093のうち中央に位置する組のユーザID表示欄2092およびユーザ映像表示枠2091に決定する(S2003)。

20

【0041】

次に、会議画面生成部209は、注目の組が画面中央に対して右および左のいずれに位置するかを示す設定値Hを右(あるいは左)に設定する(S2004)。また、カウンタ値nを1に設定する(S2005)。次に、会議画面雛形データ2090中の複数の組2093のうち、中央に位置する組から設定値Hへn番目の組があるか否かを調べる(S2006)。ない場合はS2011へ進む。一方、ある場合は、前方ユーザ検出部216から受信した前方ユーザ情報の中から、会議画面の配置位置が画面中央から設定値Hへn番目の位置である前方ユーザ情報を抽出する(S2007)。そのような前方ユーザ情報を抽出できなかった場合はS2011へ進む。一方、そのような前方ユーザ情報を抽出できた場合は、抽出した前方ユーザ情報の映像交信用IPアドレスを送信元アドレスとする映像データを、RTP処理部213から送信元アドレスと共に受信した映像データの中から検索する(S2008)。そして、S2007で抽出した前方ユーザ情報のユーザIDおよびS2008で検索した映像データの配置位置を、会議画面雛形データ2090中の複数の組2093のうち中央に位置する組から設定値Hへn番目の組のユーザID表示欄2092およびユーザ映像表示枠2091に決定する(S2009)。それから、カウンタ値nを1つインクリメントし(S2010)、S2005に戻る。

30

40

【0042】

S2011では、設定値Hが右(あるいは左)であるか否かを調べ、そうであるならば設定値Hを左(あるいは右)に変更し(S2012)、それから、S2005に戻る。一方、そうでないならば、つまり、設定値Hを右および左のそれぞれに設定して上述のS2005~S2010を実行したならば、S2013に進む。

【0043】

S2013において、会議画面生成部209は、注目の組が画面中央に対して右および

50

左のいずれに位置するかを示す設定値 J を右（あるいは左）に設定する。また、カウンタ値 m を 1 に設定する（S 2 0 1 4）。次に、会議画面雛形データ 2 0 9 0 中の設定値 J 側に位置する少なくとも 1 つの組 2 0 9 4 のうち、画面中央側から設定値 J へ m 番目の組があるか否かを調べる（S 2 0 1 5）。ない場合は S 2 0 1 9 へ進む。一方、ある場合は、前方ユーザ以外の発話ユーザの映像データであって、配置位置が前方ユーザの配置位置より設定値 J 側で画面中央から m 番目である映像データを、R T P 処理部 2 1 3 から送信元アドレスと共に受信した映像データの中から検索する（S 2 0 1 6）。

【 0 0 4 4 】

具体的には、R T P 処理部 2 1 3 から送信元アドレスと共に受信した映像データの中から、前方ユーザ検出部 2 1 6 から受信した前方ユーザ情報の映像交信用 I P アドレス以外の送信元アドレスを持つ映像データを検索する。次に、検索した映像データの各送信元アドレスを前方ユーザ検出部 2 1 6 に通知して、各送信元アドレスに対応する発話ユーザ情報を入手する。そして、入手した発話ユーザ情報の中から配置位置が前方ユーザの配置位置より設定値 J 側で画面中央から m 番目の発話ユーザ情報を特定し、該特定した発話ユーザ情報の映像交信用 I P アドレスを送信元アドレスとする映像データを、前方ユーザ以外の発話ユーザの映像データであって、配置位置が前方ユーザの配置位置より設定値 J 側で画面中央から m 番目の映像データとする。

10

【 0 0 4 5 】

次に、会議画面生成部 2 0 9 は、S 2 0 1 6 で特定した映像データと、該映像データの送信元アドレスを映像交信用 I P アドレスとする発話ユーザ情報のユーザ I D との配置位置を、会議画面雛形データ 2 0 9 0 中の設定値 J 側に位置する少なくとも 1 つの組 2 0 9 4 のうち、画面中央側から設定値 J へ m 番目の組のユーザ映像表示枠 2 0 9 1 およびユーザ I D 表示欄 2 0 9 2 に決定する（S 2 0 1 7）。それから、カウンタ値 m を 1 つインクリメントし（S 2 0 1 8）、S 2 0 1 4 に戻る。

20

【 0 0 4 6 】

S 2 0 1 9 では、設定値 J が右（あるいは左）であるか否かを調べ、そうであるならば設定値 J を左（あるいは右）に変更し（S 2 0 2 0）、それから、S 2 0 1 4 に戻る。一方、そうでないならば、つまり、設定値 J を右および左のそれぞれに設定して上述の S 2 0 1 4 ~ S 2 0 1 8 を実行したならば、会議画面生成部 2 0 9 は、会議画面雛形データ 2 0 9 0 のユーザ映像表示枠 2 0 9 1 およびユーザ I D 表示欄 2 0 9 2 のそれぞれに、S 2 0 0 9 および S 2 0 1 7 で決定したユーザ I D、映像データを配置して、図 1 0 に示すような会議画面データを生成し、映像出力部 2 0 4 を介してディスプレイ 2 2 4 に出力する（S 2 0 2 1）。なお、図 1 0 に示す会議画面データの例では、3 人の前方ユーザ（jiro, taro, hanako）の映像データが表示され、仮想空間において自ユーザに対して右側の 1 番目に位置する発話ユーザ（ichiro）の映像データが表示されている例を示している。仮想空間において自ユーザに対して左側に位置する発話ユーザは存在せず、このため、画面左側の組 2 0 9 4 にはユーザ I D および映像データが表示されていない。

30

【 0 0 4 7 】

図 5 に戻って説明を続ける。着席室情報作成部 2 1 7 は、空間モデラ 2 1 1 から受け取った自ユーザの仮想位置情報および他ユーザ各々のレコード 1 0 4 0 の仮想位置情報に基づいて、例えば図 1 1 に示すような、仮想空間における自ユーザの配置位置 2 1 6 1 および他ユーザ各々の配置位置 2 1 6 2 を示す着席情報表示データを生成する。そして、着席情報表示データを映像出力部 2 0 4 を介してディスプレイ 2 2 4 に表示する。

40

【 0 0 4 8 】

オーディオレンダラ 2 0 8 は、R T P 処理部 2 1 3 から各音声データを送信元アドレスと共に受信する。そして、受信した各音声データをバッファリングすることによって、各音声データ間で同期させる（対応付ける）。このバッファリング（プレイアウト・バッファリング）の方法については、例えば文献「Colin Perkins 著：RTP: Audio and Video for the Internet, Addison-Wesley Pub Co; 1st edition (June 11, 2003)」に記載されている。また、オーディオレンダラ 2 0 8 は、同期させた各音声データを、空間モデラ 2

50

11より受け取った他ユーザ各々のレコード1040の仮想位置情報および自ユーザの仮想位置情報により特定される、仮想空間における他ユーザ各々と自ユーザとの相対位置に基づいて立体化する。そして、オーディオレンダラ208は、2チャンネル(左チャンネルと右チャンネル)の信号データ(信号列)を、音声出力部203に接続された3Dオーディオ対応ヘッドフォン223に出力する。

【0049】

オーディオレンダラ208をより詳細に説明する。3次元オーディオ技術では、主に人の頭(以下、「人頭」)のまわりでの音響の変化の仕方(インパルス応答)を表すHRIIR(Head Related Impulse Response)と、部屋などの仮想環境によって生成される擬似的な残響とによって音の方向および距離を表現する。HRIIRは、音源と人頭との距離、および、人頭と音源との角度(水平角度および垂直角度)によって決定される。なお、オーディオレンダラ208には、予めダミーヘッドを使用して各距離および各角度毎に測定したHRIIRの数値が記憶されているものとする。また、HRIIRの数値には、左チャンネル用(ダミーヘッドの左耳で測定したもの)と、右チャンネル用(ダミーヘッドの右耳で測定したもの)とで異なる数値を使用することによって、左右、前後または上下の方向感を表現する。

10

【0050】

図12はオーディオレンダラ208の処理を説明する図である。オーディオレンダラ208は、RTP処理部213から送信元アドレスと共に送られてくる音声データ各々に関して、送信元アドレス毎に下記の計算を行う。

20

【0051】

まず、オーディオレンダラ208は、送信元アドレス毎に、RTP処理部213から当該送信元アドレスに対応付けられている音声データの信号列 $s_i[t]$ ($t=1, \dots$)を受け付ける。また、送信元アドレス毎に、空間モデラ211から受け付けた他ユーザ各々のレコード1040の中から当該送信元アドレスを音声交信用IPアドレスとするレコード1040を検索する。そして、検索したレコード1040の仮想位置情報と、空間モデラ211から受け付けた自ユーザの仮想位置情報とを、当該送信元アドレスに対応付けられている音声データの信号列 $s_i[t]$ ($t=1, \dots$)を3Dオーディオ処理に用いるパラメータに設定する(S3001)。

【0052】

次に、オーディオレンダラ208は、送信元アドレス毎に、音声データの直接音と、残響である反射音とを計算する。直接音については、パラメータ設定された位置情報を用いて、自ユーザと当該送信元アドレスを持つ他ユーザとの仮想空間における距離および角度(azimuth)を計算する(S3002)。それから、オーディオレンダラ208は、自ユーザとの距離および角度に対応するHRIIRを、予め記憶しておいたHRIIRの数値の中から特定する(S3003)。なお、オーディオレンダラ208は、予め記憶しておいたHRIIRの数値を補間することによって算出したHRIIRの数値を使用してもよい。

30

【0053】

次に、オーディオレンダラ208は、S3001で入力した信号列と、S3003で特定したHRIIRの左チャンネル用HRIIRとを使用して、畳み込み(convolution)計算を行い、左チャンネル信号を生成する(S3004)。同様に、S3001で入力した信号列と、S3003で特定したHRIIRの右チャンネル用HRIIRとを使用して、畳み込み計算を行い、右チャンネル信号を生成する(S3005)。

40

【0054】

また、反響音については、S3001でパラメータ設定された位置情報を用いて、付加すべき残響を計算する(S3006、S3007)。すなわち、オーディオレンダラ208は、仮想空間の属性による音響の変化の仕方(インパルス応答)に基づいて残響を計算する。以下、残響の計算について説明する。

【0055】

残響は初期反射(early reflection)および後期残響(late reverberation)により構

50

成される。そして、初期反射の方が後期残響より、他ユーザとの距離や部屋（仮想空間）の大きさなどに関する感覚の形成（認知）において、重要であると一般的に考えられている。実空間上の室内では、音源から直接発せられた音（直接音）が聞こえた後、数msから100msくらいの間に、条件によっては、壁、天井、床などからの数10個の初期反射を聞くことができるといわれている。部屋の形状が直方体であれば、1回の初期反射は6個だけである。しかしながら、より複雑な形状または家具などがある部屋においては、反射音の数が増え、また、壁などで複数回反射した音も聞こえる。

【0056】

初期反射の計算法としてimage source methodがあり、例えば文献「Allen, J.B. and Berkley, A., "Image Method for efficiently Simulating Small-Room Acoustics", J. Acoustical Society of America, Vol.65, No.4, pp.943-950, April 1979.」に記載されている。単純なimage source methodでは、部屋の壁、天井、床を鏡面とみなし、反射音を鏡面の反対側にある音源の像からの音として計算する。

【0057】

図13は説明を簡単にするために、天井と床を省略した2次元のimage source methodを模式的に表した図である。すなわち、中央に本来の会議室（仮想空間）2081があり、当該会議室2081には会議の参加者である自ユーザおよび他ユーザが存在する。そして、音室2081の周囲には部屋の壁2082を含む12個の鏡像が描かれている。なお、鏡像は12個である必然性はなく、これより多くすることも少なくすることもできる。

【0058】

オーディオレンダラ208は、鏡像各々の中に存在する他ユーザの各像からの音が、自ユーザ（聴取者）に直進するものとして、他ユーザの各像から自ユーザまでの距離と方向を算出する（S3006）。音の強さは距離に反比例するため、オーディオレンダラ208は、距離に従って各音量を減衰させる。ただし、壁の反射率を（0 1）とすると、壁でn回反射される音の標本には、 1^n を乗じて音量をさらに減衰させる。

【0059】

なお、反射率の値は0.6程度の値を使用する。0.6程度の値にする理由は、自ユーザが他ユーザとの距離を認識するのに十分な残響（すなわち、直接音と反射音との比）を取得するためである。また、もう1つの理由としては、の値を過大にした場合、自ユーザの方向感覚をにぶらせるからである。

【0060】

次に、オーディオレンダラ208は、他ユーザの像毎に、自ユーザとの距離および角度に対応するHRIIRを、予め記憶しておいたHRIIRの数値の中から特定する（S3007）。反射音はそれぞれ異なる方向から人頭に達するため、S3003で特定した直接音のHRIIRとは異なるHRIIRを適用する必要がある。

【0061】

なお、多数の反射音各々に、異なるHRIIRを用いて後述するたたみこみ計算（S3007、S3008）を行うと、膨大な計算が必要になる。計算量の増加を防止するため、反射音の計算には、実際の音源の方向にかかわらず正面に音源があるときのHRIIRを適用してもよい。そして、音が左右の耳に達する際の時間差（ITD: interaural time difference）と強度差（IID: interaural intensity difference）だけを計算することで、少ない計算量でHRIIRの計算を代替できる。

【0062】

次に、オーディオレンダラ208は、S3001で入力した信号列と、S3007で特定したHRIIRの左チャンネル用HRIIRとを使用して、畳み込み計算を行い、左チャンネル信号の残響を生成する（S3008）。同様に、S3001で入力した信号列と、S3007で特定したHRIIRの右チャンネル用HRIIRとを使用して、畳み込み計算を行い、右チャンネル信号の残響を生成する（S3009）。

【0063】

さて、オーディオレンダラ208は、以上のようにして全ての他ユーザ（送信元アドレ 50

ス)の左チャンネル信号を計算したならば、これらを全て加算する(S3010)。なお、左チャンネル信号は、S3004で算出した直接音と、S3008で算出した反射音とが含まれる。

【0064】

同様に、オーディオレンダラ208は、以上のようにして全ての他ユーザ(送信元アドレス)の右チャンネル信号を計算したならば、これらを全て加算する(S3011)。なお、右チャンネル信号は、S3005で算出した直接音と、S3009で算出した反射音とが含まれる。

【0065】

H R I R 計算(S3003、S3007)は、R T P パケットの1パケット分の音声データ毎に行う。しかし、畳み込み計算(S3004、S3005、S3008、S3009)では、次の1パケット分の音声データに繰り越すべき部分が生じる。このため、特定したH R I Rまたは入力された信号列を次の1パケット分の音声データに対する処理まで保持する必要がある。

【0066】

このように、オーディオレンダラ208は、R T P 処理部213から送られてきた他ユーザ各々の音声データに対して、上述の計算による音量の調節、残響や反響音の重ね合わせ、および、フィルタリング等の処理を行い、自ユーザの仮想空間内の位置において聞こえるべき音に音響効果を施す。すなわち、オーディオレンダラ208は、仮想空間の属性と他ユーザとの相対的な位置から帰結する処理によって音声定位させた立体音響を生成する。

【0067】

図5に戻って説明を続ける。S I P 制御部214は、他の会議端末2各々のユーザID、音声交信用S I P - U R Iおよび映像交信用S I P - U R Iが登録されたテーブルを保持しており、必要に応じてこのテーブルを用いて、他の会議端末2との間に、音声交信用のコネクションおよび映像交信用のコネクションを確立する。

【0068】

図14はS I P 制御部214の動作フローを説明する図である。

【0069】

会議端末2の起動時に、S I P 処理部214は、会議に参加している各会議端末2と音声交信用コネクションを確立する。まず、S I P 処理部214は、自身のテーブルに登録されている音声交信用S I P - U R Iの中から未抽出の音声交信用S I P - U R Iを抽出する(S4001)。次に、S I P 処理部214は、抽出した音声交信用S I P - U R Iを宛先とするI N V I T Eメッセージを、I P 網インターフェース部212を介してプレゼンスサーバ2に送信し、音声交信用S I P - U R Iを持つ会議端末2に対して、音声交信用コネクションの確立を試みる(S4002)。次に、S I P 処理部214は、自身のテーブルに登録されている全ての音声交信用S I P - U R Iを抽出したか否かを調べ(S4003)、抽出していない場合はS4001に戻り、抽出した場合は、起動時の音声交信用コネクションを確立処理を終了し、各種イベントの待ち状態に移行する。

【0070】

さて、S I P 処理部214は、I P 網インターフェース部212を介してI P 網3からI N V I T Eメッセージを受信すると(S4101)、該I N V I T Eメッセージの送信元(発側)の会議端末2との間でS I P に従った呼制御シーケンスを実行し、当該会議端末2との間にコネクションを確立する(S4102)。また、S I P 制御部214は、該I N V I T Eメッセージで宛先に指定されている自会議端末2のS I P - U R Iが映像交信用S I P - U R Iであるか、それとも音声交信用S I P - U R Iであるかを調べ(S4103)、映像交信用S I P - U R Iである場合は映像交信用コネクションの確立を、当該コネクションによる通話相手の会議端末2のI P アドレスと共にR T P 処理部213に通知する。一方、音声交信用S I P - U R Iである場合は音声交信用コネクションの確立を、当該コネクションによる通話相手の会議端末2のI P アドレスと共にR T P 処理部21

10

20

30

40

50

3に通知する(S4104)。

【0071】

また、SIP処理部214は、IP網インターフェース部212を介してコネクションを確立している通話相手の会議端末2からBYEメッセージを受信すると(S4201)、該通話相手の会議端末2との間でSIPに従った呼制御シーケンスを実行し、当該会議端末2との間のコネクションを解放する(S4202)。また、SIP制御部214は、解放したコネクションの種別(映像交信用コネクションおよび音声交信用コネクションのいずれか)と、該コネクションを用いた通話相手の会議端末2のIPアドレスとを、RTP処理部213に通知して、RTP処理部213に、該IPアドレスを持つ通話相手の会議端末2と指定された種別のデータを格納したRTPパケットを送受することを終了させる(S4203)。

10

【0072】

また、SIP制御部214は、前方ユーザ検出部216より前方ユーザのユーザIDを通知されると(S4301)、該ユーザIDに対応付けられて自身のテーブルに登録されている映像交信用SIP-URIを特定する。そして、特定した映像交信用SIP-URIを宛先とするINVITEメッセージを、IP網インターフェース部212を介してプレゼンスサーバ2に送信し、映像交信用SIP-URIを持つ会議端末2(前方ユーザの会議端末2)に対して、映像交信用コネクションの確立を試みる(S4302)。

【0073】

また、SIP制御部214は、発話検出部215より発話検出信号を通知されると(S4401)、自身のテーブルに登録されている映像交信用SIP-URIの中から映像交信用コネクションが確立していない映像交信用SIP-URIを特定する。そして、特定した映像交信用SIP-URI各々を宛先とするINVITEメッセージを、IP網インターフェース部212を介してプレゼンスサーバ2に送信し、該映像交信用SIP-URIを持つ会議端末2(前方ユーザ以外の他ユーザの会議端末2)に対して、映像交信用コネクションの確立を試みる(S4402)。

20

【0074】

また、SIP制御部214は、発話検出部215より発話停止信号を通知されると(S4501)、自身のテーブルに登録されている映像交信用SIP-URIの中から前方ユーザ以外の他ユーザの映像交信用SIP-URIであって、映像交信用コネクションが確立中の映像交信用SIP-URIを特定する。そして、特定した映像交信用SIP-URI各々を宛先とするBYEメッセージを、IP網インターフェース部212を介して該映像交信用SIP-URIを持つ会議端末2に送信し、該映像交信用SIP-URIを持つ会議端末2(前方ユーザ以外の他ユーザの会議端末2)との映像交信用コネクションを解放する(S4502)。

30

【0075】

上記構成のプレゼンスサーバ1には、図15に示すような、プログラムに従ってデータの加工・演算を行なうCPU301と、CPU301が直接読み書き可能なメモリ302と、ハードディスク等の外部記憶装置303と、IP網3を介して外部システムとデータ通信をするための通信装置304と、を有する一般的なコンピュータシステムを利用することができる。具体的には、サーバ、ホストコンピュータなどである。

40

【0076】

また、上記構成の会議端末2も、図15に示すような、プログラムに従ってデータの加工・演算を行なうCPU301と、CPU301が直接読み書き可能なメモリ302と、ハードディスク等の外部記憶装置303と、IP網3を介して外部システムとデータ通信をするための通信装置304と、入力装置305と、出力装置306とを有する一般的なコンピュータシステムを利用することができる。例えば、PDA(Personal Digital Assistant)、PC(Personal Computer)などのコンピュータシステムである。

【0077】

上記各装置の各機能は、メモリ302にロードまたは記憶された所定のプログラム(プ

50

レゼンサーバ1の場合はプレゼンサーバ用のプログラム、そして、会議端末2の場合は会議端末用のプログラム)を、CPU301が実行することにより実現される。

【0078】

次に、上記構成のビデオ会議システムの概略動作を説明する。

【0079】

図16は、図1に示すビデオ会議システムの概略動作を説明するための図である。ここでは、他ユーザA~Dの会議端末2が既に会議に参加中であり、後から自ユーザの会議端末2が会議に参加する場合を例にとり説明する。

【0080】

まず、自ユーザの会議端末2は、他ユーザA~Dの会議端末2の音声交信用SIP-URIを宛先とするINVITEメッセージをプレゼンサーバ1に送信する。プレゼンサーバ1は、これらのINVITEメッセージを他ユーザA~Dの会議端末2に送信する(S5001)。これにより、自ユーザの会議端末2は、他ユーザA~Dの会議端末2各々との間に音声交信用コネクションを確立する(S5002)。

10

【0081】

次に、自ユーザの会議端末2は、自ユーザの仮想位置情報をプレゼンサーバ1に送信する。これを受けて、プレゼンサーバ1は自ユーザの仮想位置情報を管理する(S5003)。また、自ユーザの会議端末2は、位置情報送信要求をプレゼンサーバ1に送信する(S5004)。これを受けて、プレゼンサーバ1は他ユーザA~Dの仮想位置情報を自ユーザの会議端末2に送信する(S5005)。

20

【0082】

次に、自ユーザの会議端末2は、前方ユーザの検出を行う(S5006)。ここでは、前方ユーザとして他ユーザA~Cが検出されたものとする。この場合、自ユーザの会議端末2は、前方ユーザA~Cの会議端末2各々の映像交信用SIP-URIを宛先とするINVITEメッセージをプレゼンサーバ1に送信する。プレゼンサーバ1は、これらのINVITEメッセージを前方ユーザA~Cの会議端末2各々に送信する(S5007)。これにより、自ユーザの会議端末2は、前方ユーザA~Cの会議端末2各々との間に映像交信用コネクションを確立する(S5008)。

【0083】

さて、自ユーザの会議端末2は、自ユーザの発話を検出すると(S5009)、前方ユーザ以外の他ユーザであって、映像交信用コネクションが未確立である他ユーザDの会議端末2の映像交信用SIP-URIを宛先とするINVITEメッセージをプレゼンサーバ1に送信する。プレゼンサーバ1は、このINVITEメッセージを他ユーザDの会議端末2に送信する(S5010)。これにより、自ユーザの会議端末2は、他ユーザDの会議端末2との間に映像交信用コネクションを確立する(S5011)。

30

【0084】

また、自ユーザの会議端末2は、自ユーザの発話停止を検出すると(S5012)、前方ユーザ以外の他ユーザであって、映像交信用コネクションが確立中である他ユーザDの会議端末2にBYEメッセージを送信して、該映像交信用コネクションを解放する(S5013)。

40

【0085】

以上、本発明の第1実施形態を説明した。本実施形態では、会議端末2のユーザが指定した当該ユーザの仮想空間における位置および向き対して前方に位置するユーザ各々の映像データを当該会議端末2に表示する。このため、ユーザ全員の映像データを会議端末2に表示する場合に比べて必要な帯域を小さくできる。また、会議端末2のユーザが、仮想空間において所望のユーザが自身の前方に位置するように自身の仮想空間における位置および向きを指定することにより、発話の有無に関わらずに、所望のユーザの様子を確認できる。したがって、本実施形態によれば、必要な帯域を低く抑えることができ、且つ、その会議のキーパーソン等の様子を容易に確認できる。

【0086】

50

<<第2実施形態>>

図17は本発明の第2実施形態が適用されたビデオ会議システムの概略構成図である。図示するように、本実施形態のビデオ会議システムは、プレゼンスサーバ1'、映像サーバ4、音声サーバ5および複数の会議端末2'がIP網3を介して相互接続されて構成されている。

【0087】

プレゼンスサーバ1'は、各会議端末2'からユーザの仮想位置情報を受信して管理する。また、映像サーバ4または音声サーバ5からの位置情報送信要求に応答して、各会議端末2'のユーザの仮想位置情報を映像サーバ4または音声サーバ5に送信する。なお、本実施形態のプレゼンスサーバ1'は、図2に示す第1実施形態のプレゼンスサーバ1からSIP処理部103を省略したものである。本実施形態のプレゼンスサーバ1'の処理フローは、図4に示す第1実施形態のプレゼンスサーバ1の処理フローからSIP処理(S1007~S1009)を省略したものと同一になる。

【0088】

映像サーバ4は、会議端末2'毎に、当該会議端末2'のユーザの映像データを受信する。また、会議端末2'毎に、当該会議端末2'のユーザ向けの会議画面データを生成し、当該会議端末2'に送信する。図18は映像サーバ4の概略図である。

【0089】

図示するように、映像サーバ4は、IP網3に接続するためのIP網インターフェース部401と、RTP処理部402と、SIP処理部403と、プレゼンスプロバイダ404と、空間モデラ405と、前方ユーザ検出部406と、発話通知受信部407と、映像分配部408と、会議端末2'毎に設けられた会議画面生成部409と、を有する。

【0090】

SIP制御部403は、IP網インターフェース部401を介して各会議端末2'との間に映像交信用コネクションを確立する。

【0091】

RTP処理部402は、会議端末2'毎に、当該会議端末2'との間で確立されている映像交信用コネクションを介して当該会議端末2'からユーザの映像データを受信し、受信した映像データを、該映像データの送信元アドレスと共に、映像分配部408に出力する。また、RTP処理部402は、会議端末2'毎に、当該会議端末2'に対応付けられた会議画面生成部409から出力された会議画面データを、当該会議端末2'との間で確立されている映像交信用コネクションを介して当該会議端末2'に送信する。

【0092】

プレゼンスプロバイダ404は、IP網インターフェース部401を介してプレゼンスサーバ1に位置情報送信要求を定期的に送信し、その応答としてプレゼンスサーバ1から各会議端末2'のユーザのレコード(仮想位置情報)1040を受信する。そして、受信した各ユーザのレコード1040を空間モデラ405に通知する。

【0093】

空間モデラ405は、プレゼンスプロバイダ404から各会議端末2'のユーザのレコード1040を受け取って保持すると共に、前方ユーザ検出部406に出力する。

【0094】

発話通知受信部407は、IP網インターフェース部404を介していずれかの会議端末2'から発話通知を受信すると、該発話通知を該発話通知の送信元アドレスと共に前方ユーザ検出部406に通知する。また、発話停止通知を受信すると、該発話停止通知を該発話停止通知の送信元アドレスと共に前方ユーザ検出部406に通知する。

【0095】

前方ユーザ検出部406は、会議端末2'毎に、空間モデラ211から受信した各ユーザのレコード1040の仮想位置情報に基づいて、仮想空間において当該会議端末2'のユーザの前方に位置する他ユーザを前方ユーザとして検出する。また、検出した前方ユーザ各々の会議画面における配置位置を決定する。そして、検出した前方ユーザ毎に、当該

10

20

30

40

50

前方ユーザのレコード1040に含まれているユーザIDおよび映像交信用IPアドレスと当該前方ユーザの会議画面における配置位置とを含む前方ユーザ情報を生成し、該前方ユーザ情報を対象となる会議端末2'のユーザIDに対応付けて映像分配部408に送信する。なお、前方ユーザの具体的な検出方法および配置位置決定方法は、第1実施形態の前方ユーザ検出部216と基本的に同様である(図6(A)参照)。

【0096】

また、前方ユーザ検出部406は、発話通知受信部407から発話通知を該発話通知の送信元アドレスと共に受け取ると、会議端末2'毎に、当該送信元アドレスが、当該会議端末2'のユーザに対する前方ユーザとして検出した何れかのユーザの映像交信用IPアドレスと一致するか否かを調べる。そして、一致しない場合に、当該送信元アドレスを持つユーザを、当該会議端末2'のユーザに対する発話ユーザとして検出し、検出した発話ユーザの会議画面における配置位置を決定する。そして、検出した発話ユーザのレコード1040に含まれているユーザIDおよび映像交信用IPアドレス(発話通知の送信元アドレス)と当該発話ユーザの会議画面における配置位置とを含む発話ユーザ情報を生成し、該発話ユーザ情報を対象となる会議端末2'のユーザIDに対応付けて映像分配部408に送信する。なお、発話ユーザの具体的な配置位置決定方法は、第1実施形態の前方ユーザ検出部216と基本的に同様である(図6(B)参照)。

10

【0097】

また、前方ユーザ検出部406は、発話通知受信部407から発話停止通知を該発話通知の送信元アドレスと共に受け取ると、映像分配部408に出力している前方ユーザ情報のうち、当該発話停止通知の送信元アドレスを映像交信用IPアドレスとする前方ユーザ情報の生成および出力を停止する。

20

【0098】

映像分配部408は、会議端末2'毎に、RTP処理部402から受け取った各ユーザの映像データの中から、当該会議端末2'に送信する会議画面に利用する映像データを抽出する。具体的には、会議端末2'毎に次の処理を行う。すなわち、前方ユーザ検出部406から受信した前方ユーザ情報および発話ユーザ情報のうち、当該会議端末2'のユーザIDに対応付けられている前方ユーザ情報および発話ユーザ情報を、当該会議端末2'の前方ユーザ情報および発話ユーザ情報として検出する。また、RTP処理部402から受け取った各ユーザの映像データのうち、当該会議端末2'の前方ユーザ情報あるいは発話ユーザ情報の映像交信用IPアドレスを送信元アドレスとする映像データを検出する。そして、検出した映像データを、当該映像データの送信元アドレスを映像交信用IPアドレスとする当該会議端末2'の前方ユーザ情報あるいは発話ユーザ情報と共に、当該会議端末2'に対応付けられた会議画面生成部409に出力する。

30

【0099】

会議画面生成部409は、図7に示すような会議画面雛形データ2090を保持している。会議画面生成部409は、この会議画面雛形データ2090を用いて、映像分配部408から受け取った各映像データを、前方ユーザの映像データについては複数の組2093のいずれかに当該前方ユーザのユーザIDと共に表示し、前方ユーザ以外の発話ユーザの映像データについては少なくとも1つの組2094のいずれかに当該発話ユーザのユーザIDと共に表示する会議画面データを生成し、RTP処理部402に出力する。

40

【0100】

なお、各会議画面生成部409の動作フローは、図8および図9に示す第1実施形態の会議画面生成部209の動作フローと基本的に同様である。但し、次の点で異なる。

【0101】

すなわち、映像データは、前方ユーザ情報あるいは発話ユーザ情報に対応付けられて映像分配部408から送られてくる。したがって、図8のS2001、S2002において、会議画面生成部209は、映像分配部408から受信した前方ユーザ情報の中から会議画面の配置位置が画面中央である前方ユーザ情報を抽出し、映像分配部408から受信した映像データの中から前記抽出した前方ユーザ情報に対応付けられている映像データを検

50

索する。また、図8のS2007、S2008において、会議画面生成部209は、映像分配部408から受信した前方ユーザ情報の中から、会議画面の配置位置が画面中央から設定値Hへn番目の位置である前方ユーザ情報を抽出し、映像分配部408から受信した映像データの中から前記抽出した前方ユーザ情報に対応付けられている映像データを検索する。また、図9のS2016において、会議画面生成部209は、映像分配部408から受信した発話ユーザ情報の中から、会議画面雛形データ2090中の設定値J側に位置する少なくとも1つの組2094のうち、画面中央側から設定値Jへm番目の組を配置位置とする発話ユーザ情報を抽出する。また、映像分配部408から受信した映像データの中から前記抽出した発話ユーザ情報に対応付けられている映像データを検索する。そして、図9のS2021において、生成した会議画面データ(図10参照)をRTP処理部402に出力する。

【0102】

音声サーバ5は、会議端末2'毎に、当該会議端末2'のユーザの音声データを受信する。また、音声サーバ5は、会議端末2'毎に、当該会議端末2'のユーザ向けの会議音声データ(3Dオーディオデータ)を生成し、当該会議端末2'に送信する。図19は音声サーバ5の概略図である。

【0103】

図示するように、音声サーバ5は、IP網3に接続するためのIP網インターフェース部501と、RTP処理部502と、SIP処理部503と、プレゼンスプロバイダ504と、空間モデラ505と、ユーザ情報生成部506と、音声分配部508と、会議端末2'毎に設けられたオーディオレンダラ509と、を有する。

【0104】

SIP制御部503は、IP網インターフェース部501を介して各会議端末2'との間に音声交信用コネクションを確立する。

【0105】

RTP処理部502は、会議端末2'毎に、当該会議端末2'との間で確立されている音声交信用コネクションを介して当該会議端末2'からユーザの音声データを受信し、受信した音声データを、該音声データの送信元アドレスと共に、音声分配部508に出力する。また、RTP処理部502は、会議端末2'毎に、当該会議端末2'に対応付けられたオーディオレンダラ509から出力された会議音声データを、当該会議端末2'との間で確立されている音声交信用コネクションを介して当該会議端末2'に送信する。

【0106】

プレゼンスプロバイダ504は、IP網インターフェース部501を介してプレゼンスサーバ1に位置情報送信要求を定期的に送信し、その応答としてプレゼンスサーバ1から各会議端末2'のユーザのレコード(仮想位置情報)1040を受信する。そして、受信した各ユーザのレコード1040を空間モデラ505に通知する。

【0107】

空間モデラ505は、プレゼンスプロバイダ504から各会議端末2'のユーザのレコード1040を受け取って保持すると共に、ユーザ情報生成部506に出力する。

【0108】

ユーザ情報生成部506は、会議端末2'毎に、空間モデラ211から受信した各ユーザのレコード1040の中から当該会議端末2'のユーザIDを含むレコード1040を特定する。そして、特定したレコード1040に含まれているユーザID、音声交信用IPアドレスおよび仮想位置情報を含む自ユーザ情報を生成し、音声分配部508に送信する。また、ユーザ情報生成部506は、前記特定したレコード1040以外のレコード1040毎に、当該レコード1040に含まれているユーザID、音声交信用IPアドレスおよび仮想位置情報を含む他ユーザ情報を生成し、前記特定したレコード1040に含まれている自ユーザのユーザIDに対応付けて音声分配部508に送信する。

【0109】

音声分配部508は、会議端末2'毎に、RTP処理部502から受け取った各ユーザ

の音声データの中から、当該会議端末2'に送信する会議音声データに利用する音声データを抽出する。具体的には、会議端末2'毎に次の処理を行う。すわわち、ユーザ情報生成部506から受信した自ユーザ情報のうち、当該会議端末2'のユーザIDを含む自ユーザ情報を検出し、これを当該会議端末2'の自ユーザ情報として、当該会議端末2'に対応付けられたオーディオレンダラ509に出力する。また、ユーザ情報生成部506から受信した他ユーザ情報のうち、当該会議端末2'のユーザIDに対応付けられている他ユーザ情報を、当該会議端末2'の他ユーザ情報として検出する。また、RTP処理部502から受け取った各ユーザの音声データのうち、当該会議端末2'の他ユーザ情報の音声交信用IPアドレスを送信元アドレスとする音声データを検出する。そして、検出した音声データを、当該音声データの送信元アドレスを音声交信用IPアドレスとする当該会議端末2'の他ユーザ情報と共に、当該会議端末2'に対応付けられたオーディオレンダラ509に出力する。

10

【0110】

オーディオレンダラ509は、音声分配部508から各音声データを他ユーザ情報と共に受信する。また、音声分配部508から自ユーザ情報を受信する。そして、受信した各音声データをバッファリングすることによって、各音声データ間で同期させる(対応付ける)。また、オーディオレンダラ509は、同期させた各音声データを、各音声データに付与された他ユーザ情報の仮想位置情報と自ユーザ情報の仮想位置情報とにより特定される、仮想空間における各他ユーザと自ユーザとの相対位置に基づいて立体化する。そして、オーディオレンダラ509は、2チャンネル(左チャンネルと右チャンネル)の信号データ(信号列)を含む会議音声データをRTP処理部502に出力する。なお、音声データの立体化の方法は、第1実施形態のオーディオレンダラ208のそれと基本的に同様である(図12および図13参照)。

20

【0111】

会議端末2'は、映像サーバ4との間に映像交信用コネクションを確立し、該コネクションを介して自ユーザの映像データを映像サーバ4に送信する。また、該コネクションを介して映像サーバ4より会議画面データを受信して表示する。図20は会議端末2'の概略構成図である。

【0112】

図示するように、会議端末2'は、音声入力部201と、映像入力部202と、音声出力部203と、映像出力部204と、操作受付部205と、オーディオエンコーダ206と、ビデオエンコーダ207と、オーディオデコーダ238と、ビデオデコーダ239と、プレゼンスプロバイダ210と、空間モデラ211と、IP網インターフェース部212と、RTP処理部233と、SIP制御部234と、発話通知部235と、着席情報作成部217と、を有する。ここで、図5に示す第1実施形態の会議端末2と同じ機能を有するものには同じ符号を付している。

30

【0113】

SIP制御部234は、IP網インターフェース部212を介して映像サーバ4との間に映像交信用コネクションを確立する。また、IP網インターフェース部212を介して音声サーバ5との間に音声交信用コネクションを確立する。

40

【0114】

RTP処理部233は、映像サーバ4との間で確立されている映像交信用コネクションを介してビデオエンコーダ207から出力された映像データを映像サーバ4に送信すると共に、該映像交信用コネクションを介して映像サーバ4から会議画面データを受信し、受信した会議画面データをビデオデコーダ239に送信する。また、音声サーバ5の間で確立されている音声交信用コネクションを介してオーディオエンコーダ206から出力された音声データを音声サーバ5に送信すると共に、該音声交信用コネクションを介して音声サーバ5から会議音声データを受信し、受信した会議音声データをオーディオデコーダ238に送信する。

【0115】

50

ビデオデコーダ 239 は、RTP 処理部 233 から受け取った会議画面データをデコードして映像信号を映像出力部 204 に出力する。オーディオデコーダ 238 は、RTP 処理部 233 から受け取った会議音声データをデコードして音声信号を音声出力部 204 に出力する。

【0116】

発話検出部 235 は、音声入力部 201 に入力する音声信号からユーザの発話の有無をモニタする。そして、発話（有音状態）を検出した場合に、IP 網インターフェース部 212 を介して映像サーバ 4 に発話通知を送信する。また、所定期間（例えば 10 秒）の無音状態を検出した場合に、発話停止通知を映像サーバ 4 に送信する。

【0117】

上記構成のプレゼンスサーバ 1'、映像サーバ 4 および音声サーバ 5 も、第 1 実施形態のプレゼンスサーバ 1 と同様に、図 15 に示すような、プログラムに従ってデータの加工・演算を行なう CPU 301 と、CPU 301 が直接読み書き可能なメモリ 302 と、ハードディスク等の外部記憶装置 303 と、IP 網 3 を介して外部システムとデータ通信をするための通信装置 304 と、を有する一般的なコンピュータシステムを利用することができる。具体的には、サーバ、ホストコンピュータなどである。

【0118】

また、上記構成の会議端末 2' も、第 1 実施形態の会議端末 2 と同様に、図 15 に示すような、プログラムに従ってデータの加工・演算を行なう CPU 301 と、CPU 301 が直接読み書き可能なメモリ 302 と、ハードディスク等の外部記憶装置 303 と、IP 網 3 を介して外部システムとデータ通信をするための通信装置 304 と、入力装置 305 と、出力装置 306 とを有する一般的なコンピュータシステムを利用することができる。例えば、PDA、PC などのコンピュータシステムである。

【0119】

次に、上記構成のビデオ会議システムの概略動作を説明する。

【0120】

図 21 は図 17 に示すビデオ会議システムの概略動作を説明するための図である。ここで、ユーザ E ~ G の会議端末 2' 各々は、音声サーバ 5 および映像サーバ 4 のそれぞれとの間に音声交信用コネクション、映像交信用コネクションを確立しているものとする。

【0121】

さて、ユーザ E ~ G の会議端末 2' 各々は、定期的に、自ユーザの仮想位置情報をプレゼンスサーバ 1' に送信する（S6001）。これを受けて、プレゼンスサーバ 1' はユーザ E ~ G 各々の仮想位置情報を管理する。

【0122】

また、音声サーバ 5 は、定期的に、位置情報送信要求をプレゼンスサーバ 1' に送信し（S6002）、その応答として、プレゼンスサーバ 1' から全てのユーザ E ~ G の仮想位置情報を入手する（S6003）。

【0123】

また、映像サーバ 4 は、定期的に、位置情報送信要求をプレゼンスサーバ 1' に送信し（S6004）、その応答として、プレゼンスサーバ 1' から全てのユーザ E ~ G の仮想位置情報を入手する（S6005）。

【0124】

また、ユーザ E ~ G の会議端末 2' 各々は、音声サーバ 5 との間に確立されている音声交信用コネクションを介して自ユーザの音声データを音声サーバ 5 に送信する（S6006）。音声サーバ 5 は、ユーザ E ~ G の会議端末 2' から送られてきた音声データを用いて、会議端末 2' 毎に、当該会議端末 2 のユーザ E ~ G 向けの会議音声データを生成し、当該会議端末 2 との間に確立されている音声交信用コネクションを用いて当該会議端末に生成した会議音声データを送信する（S6007）。

【0125】

また、ユーザ E ~ G の会議端末 2' 各々は、映像サーバ 4 との間に確立されている映像

10

20

30

40

50

交信用コネクションを介して自ユーザの映像データを映像サーバ４に送信する（Ｓ６００８）。映像サーバ４は、ユーザＥ～Ｇの会議端末２´から送られてきた映像データを用いて、会議端末２´毎に、当該会議端末２のユーザＥ～Ｇ向けの会議画面データを生成し、当該会議端末２との間に確立されている映像交信用コネクションを用いて当該会議端末に生成した会議画面データを送信する（Ｓ６００９）。

【０１２６】

また、ユーザＥ～Ｇの会議端末２´各々は、自ユーザの発話を検出すると（Ｓ６０１０）、発話検出通知を映像サーバ４に送信する（Ｓ６０１１）。これを受けて、映像サーバ４は、発話検出通知の送信元の映像データを必要に応じて会議画面データに含める。

【０１２７】

以上、本発明の第２実施形態を説明した。本実施形態では、各会議端末２´は映像サーバ４との間にのみ映像交信用コネクションを確立すればよいので、ＩＰ網３のネットワークトラフィックを低減できる。また、会議端末２´のユーザが、仮想空間において所望のユーザが自身の前方に位置するように自身の仮想空間における位置および向きを指定することにより、発話の有無に関わらずに、所望のユーザの様子を確認できる。したがって、本実施形態によれば、必要な帯域を低く抑えることができ、且つ、その会議のキーパーソン等の様子を容易に確認できる。

【０１２８】

なお、本発明は、上記の各実施形態に限定されるものではなく、その要旨の範囲内で数々の変形が可能である。

【０１２９】

例えば、上記の各実施形態では、コネクションの確立にＳＩＰを利用する場合を例にとり説明した。しかし、本発明はこれに限定されない。例えばＨ．３２３等のＳＩＰ以外の呼制御プロトコルを利用してもよい。なお、上記の第２実施形態のように、会議端末２´と映像サーバ４、音声サーバ５との間で常時通信を行なうことを前提する場合は、呼制御プロトコルに従った呼制御シーケンスを省略できる。

【０１３０】

また、上記の第１実施形態に上記の第２実施形態の音声サーバ５を導入し、各会議端末２と音声サーバ５との間に音声交信用コネクションを確立して、会議端末２のオーディオレンダラ２０８による会議音声データ生成処理を省略してもよい。あるいは、上記の第１実施形態に上記の第２実施形態の映像サーバ４を導入し、各会議端末２と映像サーバ４との間に映像交信用コネクションを確立して、会議端末２の会議画面生成部２０９による会議画面データ生成処理を省略してもよい。

【図面の簡単な説明】

【０１３１】

【図１】図１は本発明の第１実施形態が適用されたビデオ会議システムの概略構成図である。

【図２】図２はプレゼンスサーバ１の概略構成図である。

【図３】図３は位置情報記憶部１０４の登録内容を模式的に示した図である。

【図４】図４はプレゼンスサーバ１の動作フローを説明する図である。

【図５】図５は会議端末３の概略構成図である。

【図６】図６（Ａ）は前方ユーザ検出部２１６による前方ユーザの検出および配置位置決定を説明するための図であり、図６（Ｂ）は前方ユーザ検出部２１６による発話ユーザの配置位置決定を説明するための図である。

【図７】図７は会議画面雛形データ２０９０の表示例を示す図である。

【図８】図８は会議画面生成部２０９の動作フローを説明する図である。

【図９】図９は会議画面生成部２０９の動作フローを説明する図である。

【図１０】図１０は会議画面データの表示例を示す図である。

【図１１】図１１は着席情報表示データの表示例を示す図である。

【図１２】図１２はオーディオレンダ２０８の処理を説明する図である。

10

20

30

40

50

【図 1 3】図 1 3 は天井と床を省略した 2 次元の image source method を模式的に表した図である。

【図 1 4】図 1 4 は SIP 制御部 2 1 4 の動作フローを説明する図である。

【図 1 5】図 1 5 はビデオ会議システムを構成する各装置のハードウェア構成例を示す図である。

【図 1 6】図 1 6 は図 1 に示すビデオ会議システムの概略動作を説明するための図である。

【図 1 7】図 1 7 は本発明の第 2 実施形態が適用されたビデオ会議システムの概略構成図である。

【図 1 8】図 1 8 は映像サーバ 4 の概略図である。

10

【図 1 9】図 1 9 は音声サーバ 5 の概略図である。

【図 2 0】図 2 0 は会議端末 2 ' の概略構成図である。

【図 2 1】図 2 1 は図 1 7 に示すビデオ会議システムの概略動作を説明するための図である。

【符号の説明】

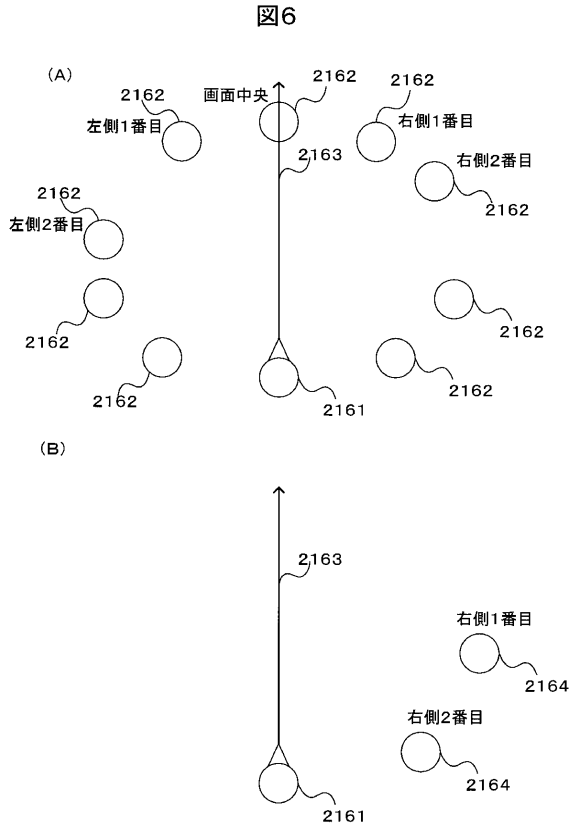
【0 1 3 2】

1、1 ' ... プレゼンスサーバ、2、2 ' ... 会議端末、3 ... IP 網、4 ... 映像サーバ、5 ... 音声サーバ、1 0 1 ... IP 網インターフェース部、1 0 2 ... 位置情報管理部、1 0 3 ... SIP サーバ処理部、1 0 4 ... 位置情報記憶部、2 0 1 ... 音声入力部、2 0 2 ... 映像入力部、2 0 3 ... 音声出力部、2 0 4 ... 映像出力部、2 0 5 ... 操作受付部、2 0 6 ... オーディオエンコーダ、2 0 7 ... ビデオエンコーダ、2 0 8 ... オーディオレンダラ、2 0 9 ... 会議画面生成部、2 1 0 ... プレゼンスプロバイダ、2 1 1 ... 空間モデラ、2 1 2 ... IP 網インターフェース部、2 1 3 ... RTP 処理部、2 1 4 ... SIP 制御部、2 1 5 ... 発話検出部、2 1 6 ... 前方ユーザ検出部、2 1 7 ... 着席情報生成部、2 3 3 ... RTP 処理部、2 3 4 ... SIP 制御部、2 3 8 ... オーディオデコーダ、2 3 9 ... ビデオデコーダ、4 0 1 ... IP 網インターフェース部、4 0 2 ... RTP 処理部、4 0 3 ... SIP 制御部、4 0 4 ... プレゼンスプロバイダ、4 0 5 ... 空間モデラ、4 0 6 ... 前方ユーザ検出部、4 0 8 ... 映像分配部、4 0 9 ... 会議画面生成部、5 0 1 ... IP 網インターフェース部、5 0 2 ... RTP 処理部、5 0 3 ... SIP 制御部、5 0 4 ... プレゼンスプロバイダ、5 0 5 ... 空間モデラ、5 0 6 ... ユーザ情報生成部、5 0 8 ... 音声分配部、5 0 9 ... オーディオレンダラ

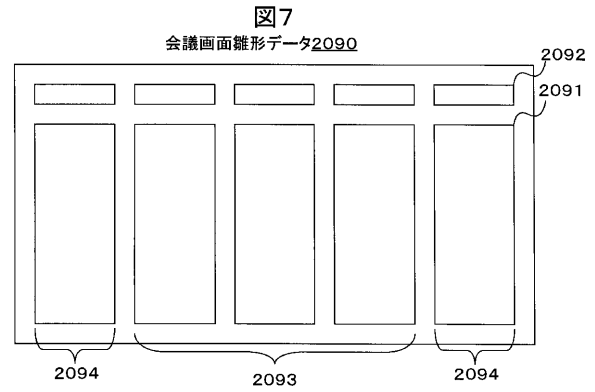
20

30

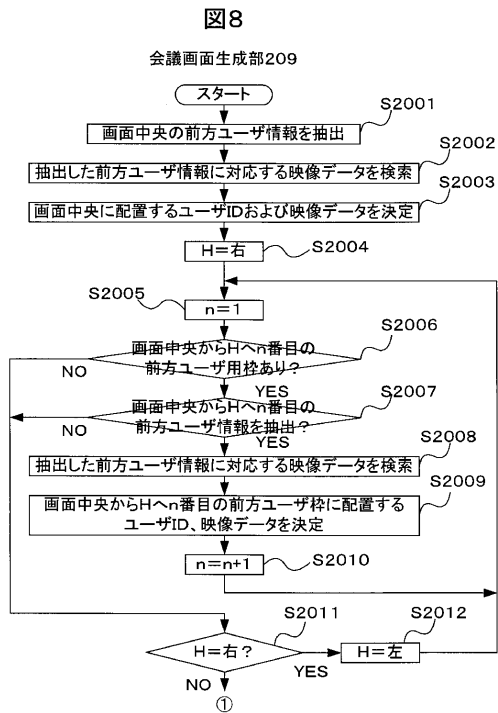
【 図 6 】



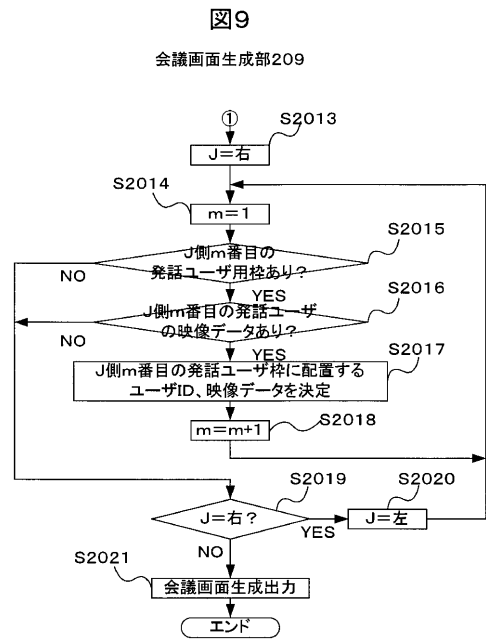
【 図 7 】



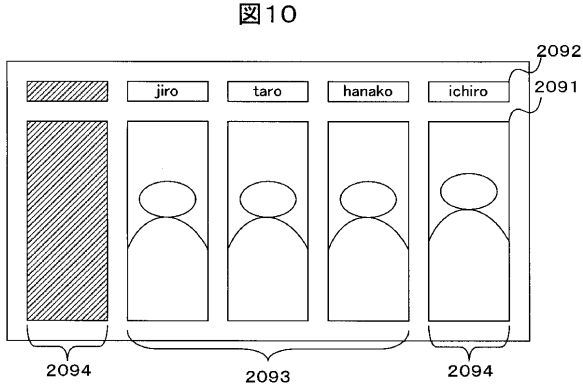
【 図 8 】



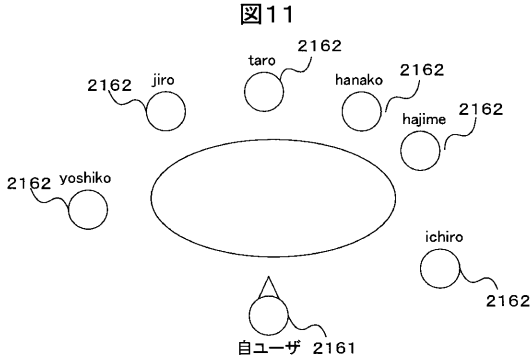
【 図 9 】



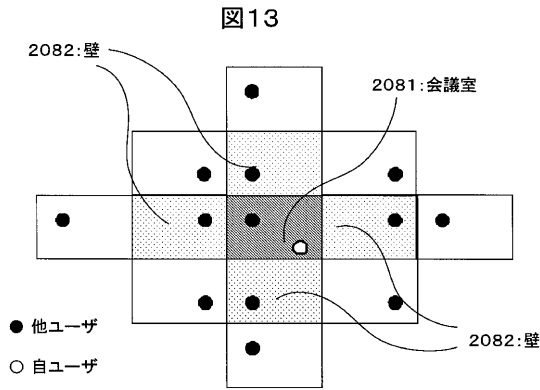
【図10】



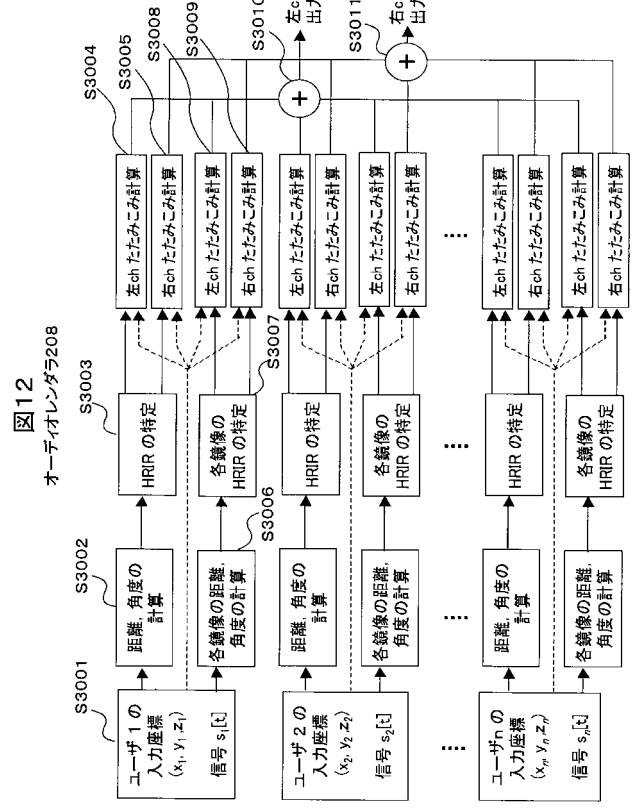
【図11】



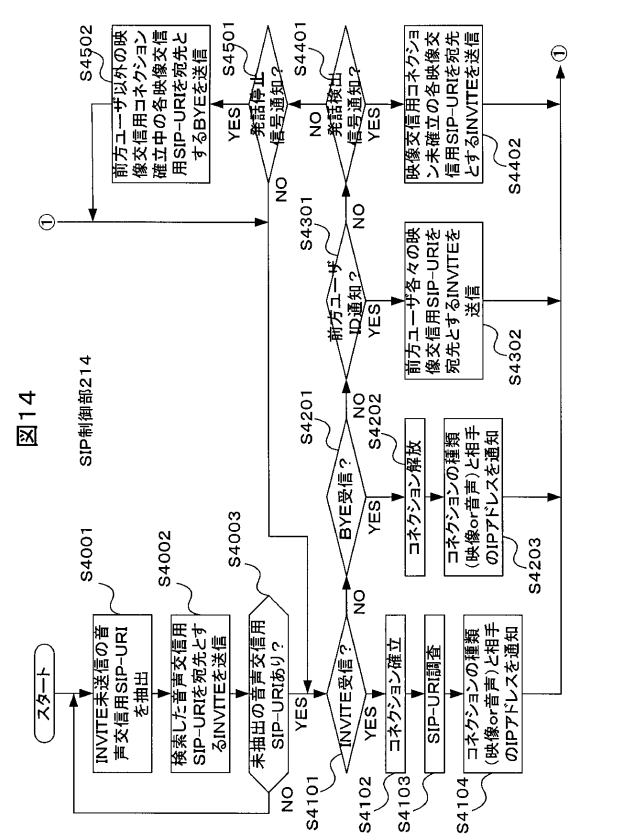
【図13】



【図12】



【図14】



【 図 1 5 】

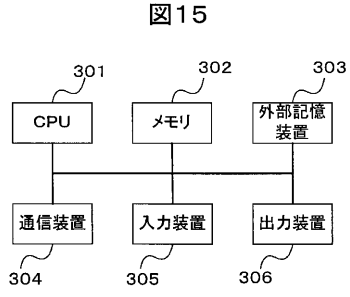


図 15

【 図 1 6 】

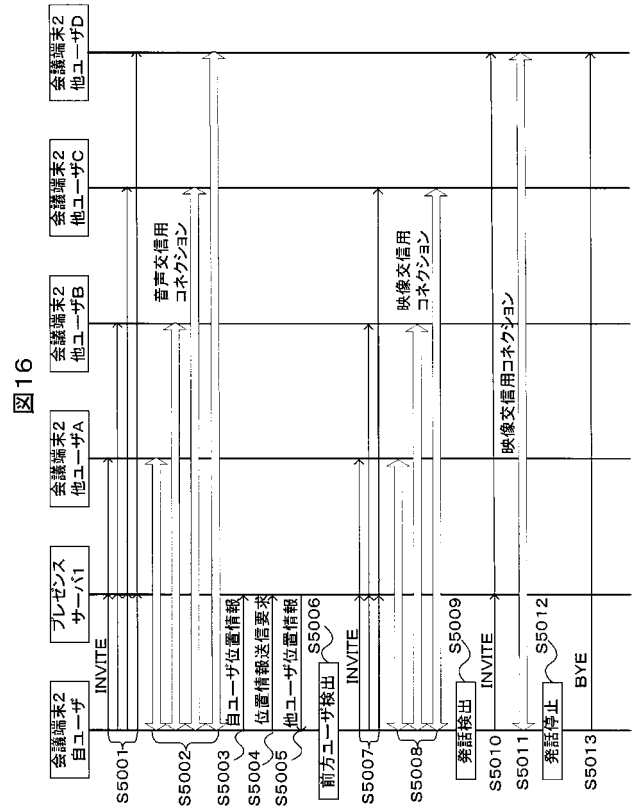


図 16

【 図 1 7 】

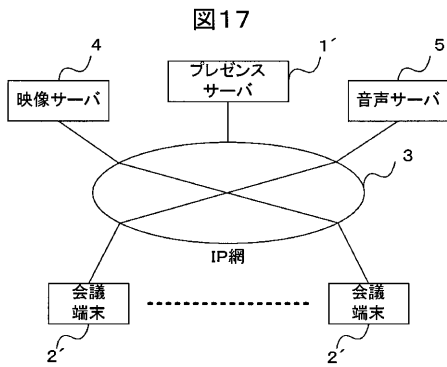


図 17

【 図 1 8 】

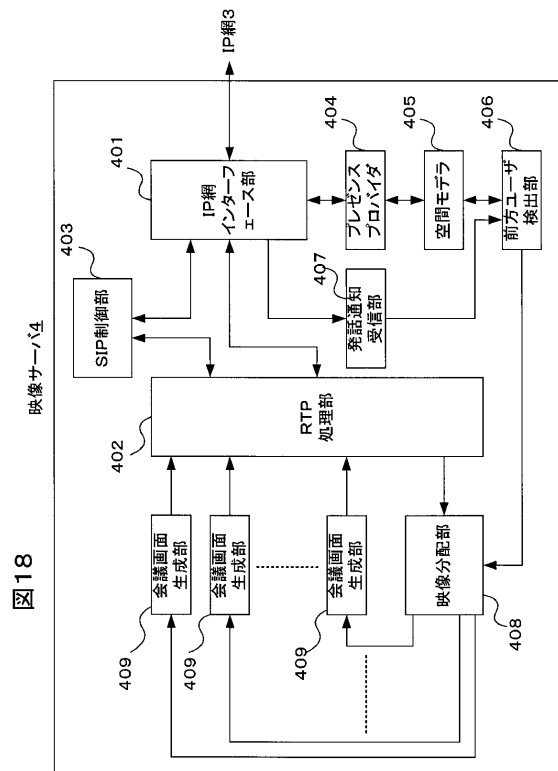
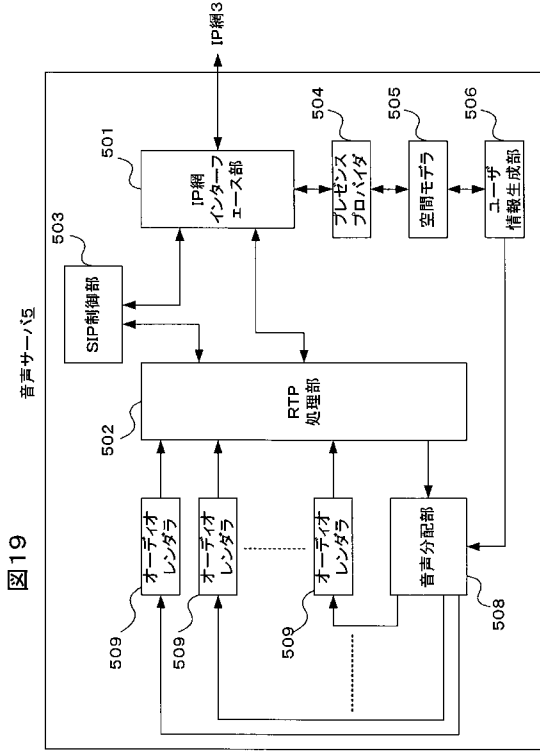
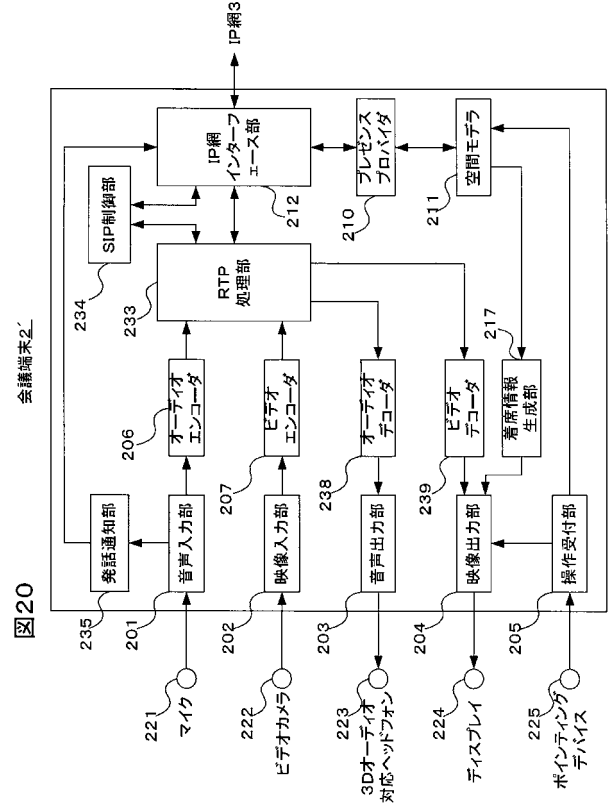


図 18

【 図 1 9 】



【 図 2 0 】



【 図 2 1 】

